

КУРС ВЫСШЕЙ МАТЕМАТИКИ



А. В. ИГНАТЬЕВА, Т. И. КРАСНОЩЕКОВА,
В. Ф. СМИРНОВ

КУРС ВЫСШЕЙ МАТЕМАТИКИ

*ПОД РЕДАКЦИЕЙ
ПРОФ. П. И. РОМАНОВСКОГО*

*Допущено
Министерством высшего и среднего
специального образования СССР
в качестве учебного пособия
для студентов высших
технических учебных
заведений*



ИЗДАТЕЛЬСТВО «ВЫСШАЯ ШКОЛА»
МОСКВА — 1964

Переплет художника В. П. Заикина

ОГЛАВЛЕНИЕ

Глава I. Аналитическая геометрия на плоскости

Раздел 1. Метод координат	11
§ 1. Направленные отрезки	11
§ 2. Координаты на прямой линии	12
§ 3. Прямоугольная декартова система координат на плоскости	15
§ 4. Простейшие задачи аналитической геометрии на плоскости	17
§ 5. Элементы теории проекций	20
§ 6. Полярная система координат	22
§ 7. Связь между прямоугольными декартовыми и полярными координатами	23
§ 8. Преобразование координат	24
Раздел 2. Линии и их уравнения	27
§ 9. Линии и их уравнения	27
Раздел 3. Прямая линия	35
§ 10. Угловой коэффициент	35
§ 11. Уравнение прямой линии с заданным угловым коэффициентом	36
§ 12. Общее уравнение прямой и его исследование	37
§ 13. Уравнение прямой, проходящей через данную точку в данном направлении; уравнение прямой, проходящей через две данные точки; уравнение прямой в отрезках на осях	38
§ 14. Угол между двумя прямыми; условия параллельности и перпендикулярности двух прямых	39
§ 15. Пересечение двух прямых	41
§ 16. Нормальное уравнение прямой	42
§ 17. Расстояние от точки до прямой	43
§ 18. Уравнение пучка прямых	44
Раздел 4. Кривые второго порядка	48
§ 19. Окружность и ее уравнение	48
§ 20. Эллипс и его уравнение	49
§ 21. Исследование уравнения эллипса	50
§ 22. Эксцентриситет и фокальные радиусы эллипса	52
§ 23. Гипербола и ее уравнение	52
§ 24. Исследование уравнения гиперболы	54
§ 25. Эксцентриситет и фокальные радиусы гиперболы	57
§ 26. Парабола и ее уравнение	58
§ 27. Исследование уравнения параболы	58
§ 28. Общее свойство кривых второго порядка — эллипса, гиперболы и параболы	59
Раздел 5. Общее уравнение линий второго порядка	64
§ 29. Уравнения кривых второго порядка с осями симметрии, параллельными координатным осям	64
§ 30. Исследование общего уравнения линий второго порядка, не содержащего члена с произведением текущих координат	65
§ 31. Упрощение общего уравнения линий второго порядка	68

Раздел 6. Определители и их применение к решению систем линейных уравнений	69
§ 32. Определители второго порядка и их свойства.	69
§ 33. Определители третьего порядка и их свойства.	71
§ 34. Применение определителей к исследованию и решению систем линейных уравнений.	75
§ 35. Применение определителей к решению некоторых задач аналитической геометрии на плоскости	83
§ 36. Определители высшего порядка и их свойства.	85
§ 37. Некоторые сведения о системах линейных уравнений. . .	91

Глава II. Аналитическая геометрия в пространстве

Раздел 1. Декартова прямоугольная система координат в пространстве	94
Раздел 2. Элементы векторной алгебры	96
§ 1. Векторные и скалярные величины.	96
§ 2. Линейные операции над векторами.	97
§ 3. Основные теоремы о проекциях векторов.	99
§ 4. Координаты и компоненты векторов.	99
§ 5. Направляющие косинусы вектора.	101
§ 6. Линейные операции над векторами, заданными в координатной форме	102
§ 7. Скалярное произведение векторов.	102
§ 8. Скалярное произведение векторов в координатной форме	104
§ 9. Векторное произведение.	105
§ 10. Векторное произведение векторов в координатной форме	106
§ 11. Смешанное произведение трех векторов.	108
Раздел 3. Простейшие задачи аналитической геометрии в пространстве	111
§ 12. Расстояние между двумя точками в пространстве	111
§ 13. Деление отрезка в данном отношении	111
Раздел 4. Поверхности и линии в пространстве и их уравнения	112
Раздел 5. Плоскость	114
§ 14. Нормальное уравнение плоскости.	114
§ 15. Общее уравнение плоскости.	115
§ 16. Исследование общего уравнения плоскости.	116
§ 17. Связка плоскостей и ее приложения.	117
§ 18. Расстояние от точки до плоскости.	118
§ 19. Угол между двумя плоскостями.	120
Раздел 6. Прямая линия в пространстве	121
§ 20. Параметрические и канонические уравнения прямой линии в пространстве.	121
§ 21. Общие уравнения прямой в пространстве.	123
§ 22. Угол между двумя прямыми.	125
Раздел 7. Прямая и плоскость	126
§ 23. Угол между прямой и плоскостью.	126
§ 24. Точка пересечения прямой с плоскостью.	128
§ 25. Пучок плоскостей.	130
Раздел 8. Поверхности второго порядка	131
§ 26. Цилиндрические поверхности.	132
§ 27. Конические поверхности.	133
§ 28. Линейчатые поверхности.	134
§ 29. Поверхности вращения.	134
§ 30. Метод сечений при исследовании формы поверхностей. .	135

Глава III. Введение в анализ

Раздел 1. Функции	141
§ 1. Числа.	141
§ 2. Функции и их графики.	143
§ 3. Некоторые замечания о построении графиков функций. .	154
Раздел 2. Числовые последовательности	158
§ 4. Основные понятия	158
§ 5. Предел последовательности.	159
§ 6. Бесконечный предел последовательности.	162
§ 7. Бесконечно малые.	164
§ 8. Теоремы о пределах.	166

§ 9. О переходе к пределу в неравенствах	169
§ 10. Монотонные последовательности	170
§ 11. Число e	172
Раздел 3. Предел функции	174
§ 12. Определение предела функции	174
§ 13. Основные теоремы о пределах функций	177
§ 14. Бесконечно малые функции и их свойства	179
§ 15. Бесконечно большие функции и их связь с бесконечно малыми	180
§ 16. Односторонние пределы функций	182
§ 17. Некоторые замечательные пределы	183
§ 18. Сравнение бесконечно малых	185
Раздел 4. Непрерывность функции	187
§ 19. Основные определения	187
§ 20. Непрерывность некоторых элементарных функций	190
§ 21. Точки разрыва и их классификация	191
§ 22. Свойства функций, непрерывных на отрезке	192
<i>Глава IV. Дифференциальное исчисление функций одной переменной</i>	
Раздел 1. Производная функции	198
§ 1. Приращение аргумента и функции	198
§ 2. Задача о проведении касательной к кривой	198
§ 3. Определение производной функции в точке. Геометрический и механический смысл производной	200
§ 4. Производные некоторых элементарных функций	204
§ 5. Основные правила дифференцирования	205
§ 6. Производная логарифмической функции	208
§ 7. Производные степенных и показательных функций	209
§ 8. Производные обратных функций	210
§ 9. Таблица формул дифференцирования	212
§ 10. Замечания о дифференцировании неявных функций	212
Раздел 2. Дифференциал. Производные и дифференциалы высших порядков	213
§ 11. Дифференциал функции	213
§ 12. Производные и дифференциалы высших порядков	216
Раздел 3. Теоремы о среднем	219
§ 13. Теорема Ролля. Теоремы Лагранжа и Коши	219
§ 14. Раскрытие неопределенностей	223
§ 15. Формула Тейлора	226
Раздел 4. Исследование функций и построение графиков	231
§ 16. Возрастание и убывание функций	231
§ 17. Экстремумы функций	234
§ 18. О наибольшем и наименьшем значениях функции	238
§ 19. Выпуклость и вогнутость. Точки перегиба	240
§ 20. Асимптоты плоских кривых	243
§ 21. Построение графиков функций по критическим точкам и асимптотам	245
§ 22. Параметрические уравнения кривых	248
§ 23. Производные функций, заданных параметрически	250
§ 24. Уравнения касательной прямой и нормальной плоскости к пространственной кривой, заданной параметрически	251
§ 25. Кривизна. Радиус кривизны	252
§ 26. Соприкасающийся круг. Центр кривизны	256
§ 27. Круг кривизны кривой. Эволюта и эвольвента	258
§ 28. Приближенное решение уравнений	262
§ 29. Интерполирование функций	265
<i>Глава V. Неопределенный интеграл</i>	
Раздел 1. Первообразная функция и неопределенный интеграл	273
§ 1. Простейшие свойства неопределенного интеграла	275
§ 2. Непосредственное интегрирование	277
§ 3. Замена переменной в неопределенном интеграле	278
§ 4. Интегрирование по частям	279
Раздел 2. Комплексные числа	281
§ 5. Комплексные числа и действия над ними	281
§ 6. Тригонометрическая форма комплексного числа	283
§ 7. Действия над комплексными числами в тригонометрической форме	284

Раздел 3. Общие приемы интегрирования	287
§ 8. Интегрирование элементарных дробей	287
§ 9. Интегрирование рациональных функций	290
§ 10. Интегрирование тригонометрических выражений	296
§ 11. Интегрирование иррациональных функций	300
§ 12. Некоторые типы интегралов, берущихся методом интегрирования по частям	305
§ 13. Примеры некоторых интегралов, не выражающихся через элементарные функции	306

Глава VI. Определенный интеграл

Раздел 1. Определенный интеграл и его свойства	308
§ 1. Задачи, приводящие к понятию определенного интеграла	308
§ 2. Интегральная сумма. Определенный интеграл	310
§ 3. Условия существования определенного интеграла	311
§ 4. Дополнения к понятию определенного интеграла	316
§ 5. Свойства определенного интеграла	316
§ 6. Определенный интеграл с переменным верхним пределом	322
§ 7. Формула Ньютона — Лейбница	324
§ 8. Замена переменной в определенном интеграле	326
§ 9. Формула интегрирования по частям в определенном интеграле	328

Раздел 2. Геометрические приложения определенного интеграла	328
§ 10. Площадь плоской фигуры	328
§ 11. Объем тела	333
§ 12. Длина дуги кривой	336
§ 13. Площадь поверхности вращения	342

Раздел 3. Приближенные методы вычисления определенных интегралов	345
§ 14. Приближенные методы вычисления определенных интегралов	345

Раздел 4. Обобщение понятия определенного интеграла	352
§ 15. Несобственные интегралы	352

Глава VII. Дифференциальное исчисление функций нескольких переменных

Раздел 1. Функции нескольких переменных. Основные понятия	360
§ 1. Функции нескольких переменных	360
§ 2. Понятие предела для функции двух и более переменных	265
§ 3. Непрерывность функций нескольких переменных	367

Раздел 2. Производные и дифференциалы функций нескольких переменных	370
§ 4. Частные производные	372
§ 5. Дифференцируемые функции нескольких переменных	372
§ 6. Дифференциал	375
§ 7. Касательная плоскость и нормаль к поверхности, заданной уравнением $z = f(x, y)$. Геометрический смысл полного дифференциала для функции двух переменных	376
§ 8. Производные сложных функций	378
§ 9. Инвариантная форма записи дифференциала	380
§ 10. Производная по направлению	381
§ 11. Неявные функции	383
§ 12. Частные производные высших порядков	390
§ 13. Дифференциалы высших порядков	392
§ 14. Формула Тейлора для функции двух переменных	394

Раздел 3. Экстремумы функций нескольких переменных	395
§ 15. Экстремумы функций нескольких переменных	395
§ 16. Наибольшее и наименьшее значения функции	400
§ 17. Условный экстремум	401

Глава VIII. Обыкновенные дифференциальные уравнения

Раздел 1. Дифференциальные уравнения первого порядка	405
§ 1. Задачи, приводящие к дифференциальным уравнениям. Основные определения	405

§ 2. Дифференциальные уравнения первого порядка, разрешенные относительно производной, и их геометрический смысл. Основные понятия	408
§ 3. Некоторые дифференциальные уравнения первого порядка, интегрируемые в квадратурах	412
§ 4. Уравнения первого порядка, не разрешенные относительно производной	421
§ 5. Особые решения	425
§ 6. Составление дифференциального уравнения однопараметрического семейства кривых	430
§ 7. Задача об изогональных траекториях	431
§ 8. Приближенные методы интегрирования дифференциальных уравнений первого порядка	433
§ 9. Применение дифференциальных уравнений первого порядка к решению некоторых задач механики и физики	443
Раздел 2. Общие сведения о дифференциальных уравнениях высшего порядка	446
§ 10. Дифференциальные уравнения высшего порядка. Общие понятия	446
§ 11. Уравнения высшего порядка, допускающие понижение порядка	451
Раздел 3. Общая теория линейных дифференциальных уравнений высшего порядка	454
§ 12. Общие сведения о линейных дифференциальных уравнениях высшего порядка	454
§ 13. Линейные однородные дифференциальные уравнения с произвольными коэффициентами	455
§ 14. Линейная зависимость и линейная независимость функций. Определитель Вронского и его свойства	456
§ 15. Структура общего решения линейного однородного уравнения	461
§ 16. Нахождение общего решения линейного однородного дифференциального уравнения второго порядка	463
§ 17. Линейные однородные уравнения высшего порядка с постоянными коэффициентами	464
§ 18. Линейные неоднородные дифференциальные уравнения с произвольными коэффициентами. Структура общего решения	470
§ 19. Линейные неоднородные уравнения высшего порядка с постоянными коэффициентами	475
§ 20. Применение теории линейных дифференциальных уравнений к исследованию некоторых колебаний	480
Раздел 4. Некоторые сведения о нормальных системах дифференциальных уравнений	486
§ 21. Основные понятия	486
§ 22. Системы линейных дифференциальных уравнений	491
 <i>Глава IX. Интегральное исчисление функций многих переменных</i>	
Раздел 1. Интегралы, зависящие от параметров	495
§ 1. Интегралы, зависящие от параметров	495
Раздел 2. Криволинейные интегралы первого и второго рода	501
§ 2. Криволинейные интегралы первого рода (по длине дуги)	501
§ 3. Криволинейные интегралы второго рода	505
§ 4. Условия независимости криволинейного интеграла (второго рода) от формы пути интегрирования	513
Раздел 3. Двойные интегралы	520
§ 5. Задача об объеме цилиндрического тела	520
§ 6. Определение двойного интеграла	522
§ 7. Условия существования двойного интеграла и его свойства	523
§ 8. Геометрический смысл двойного интеграла	524
§ 9. Теорема о среднем	526
§ 10. Вычисление двойных интегралов	526
§ 11. Некоторые геометрические и механические приложения двойных интегралов	532
§ 12. Формула Грина — Остроградского	537
§ 13. Замена переменных в двойном интеграле	540

Раздел 4. Тройные интегралы	548
§ 14. Задача о массе тела	548
§ 15. Определение тройного интеграла и его свойства	549
§ 16. Вычисление тройных интегралов	551
§ 17. Некоторые механические приложения тройных интегралов	554
§ 18. Замена переменных в тройном интеграле	555
Раздел 5. Поверхностные интегралы первого рода (по площади поверхности)	559
§ 19. Определение поверхностного интеграла первого рода	559
§ 20. Условия существования и вычисление поверхностных интегралов первого рода	561
§ 21. Некоторые приложения поверхностных интегралов первого рода	563
Раздел 6. Поверхностные интегралы второго рода (по координатам)	564
§ 22. Двухсторонние поверхности. Ориентация поверхности	564
§ 23. Определения поверхностных интегралов второго рода (по координатам) и их простейшие свойства	566
§ 24. Существование и вычисление поверхностных интегралов второго рода	569
§ 25. Связь между поверхностными интегралами первого и второго рода	573
§ 26. Формула Гаусса — Остроградского	574
§ 27. Формула Стокса	577
<i>Глава X. Ряды</i>	
Раздел 1. Числовые ряды	580
§ 1. Общие сведения	580
§ 2. Ряды с неотрицательными членами	584
§ 3. Несобственные интегралы	588
§ 4. Знакопеременные ряды	597
§ 5. Ряды с комплексными членами	603
Раздел 2. Функциональные последовательности и ряды	605
§ 6. Сходимость и равномерная сходимость	605
§ 7. Интегрирование и дифференцирование функциональных последовательностей и рядов	613
§ 8. Понятие о функциональных рядах в комплексной области	618
Раздел 3. Степенные ряды	619
§ 9. Интервал сходимости. Характер сходимости степенного ряда	619
§ 10. Интегрирование и дифференцирование степенных рядов	624
§ 11. Ряд Тейлора. Основные разложения	627
§ 12. Операции над степенными рядами	632
§ 13. Приложения степенных рядов	635
Раздел 4. Некоторые элементарные функции комплексной переменной	640
§ 14. Показательная, тригонометрические и гиперболические функции	640
§ 15. Некоторые многозначные функции комплексного аргумента	646
Раздел 5. Ряды и интеграл Фурье	650
§ 16. Основные сведения. Тригонометрические ряды Фурье для функций с периодом 2π	650
§ 17. Тригонометрические ряды Фурье для функций любого периода	660
§ 18. Ряд Фурье по любой ортогональной системе функций. Минимальное свойство коэффициентов Фурье	663
§ 19. Комплексная форма ряда Фурье	670
§ 20. Задача о колебании струны	673
§ 21. Интеграл Фурье	677

ПРЕДИСЛОВИЕ РЕДАКТОРА

Математика, по определению Энгельса, есть наука, изучающая пространственные формы и количественные отношения реального мира в самом широком понимании этих слов.

Школьный курс математики (так называемая элементарная математика), как правило, имеет дело с постоянными величинами и неизменяемыми объектами. Курс математики (так называемой высшей математики), который предстоит изучать будущему инженеру, в основном имеет дело с переменными величинами, с процессами их изменения и их взаимозависимостями.

Основной курс высшей математики в высшем техническом учебном заведении по установившейся традиции складывается из элементов аналитической геометрии и математического анализа.

Предметом аналитической геометрии является исследование геометрических форм аналитическими методами. Положение геометрических объектов изучается аналитическими средствами (положение точек определяется координатами, положения линий и поверхностей определяются уравнениями).

В математическом анализе в основном изучаются переменные величины в их взаимозависимости, причем ведущую роль играют предельные переходы (основное понятие дифференциального исчисления — производная — появляется как предел некоторого отношения; основное понятие интегрального исчисления — интеграл — появляется как предел некоторой суммы).

Предлагаемое учебное пособие по основному курсу высшей математики для высших технических учебных заведений машиностроительного профиля составлено сотрудниками кафедры математики Московского авиационного института имени Серго Орджоникидзе: канд. физ.-мат. наук доцентом А. В. Игнатъевой, канд. физ.-мат. наук доцентом Т. И. Краснощековой и канд. пед. наук доцентом В. Ф. Смирновым.

А. В. Игнатъевой написаны гл. VI—IX, а также разделы 1, 2, 6 гл. I и § 2 гл. III; Т. И. Краснощековой — гл. III (кроме

§ 2), IV, X; В. Ф. Смирновым — разделы 3 — 5 гл. I, гл. II и V.

Главы I—II, посвященные аналитической геометрии, следует в основном рассматривать не как курс аналитической геометрии, а как своего рода геометрическое введение к изучению курса математического анализа (главы III—X).

В настоящем учебном пособии излагаются только программные вопросы основного втузовского курса высшей математики. Различные дополнительные разделы математики, ныне входящие в программу отдельных втузов, сюда не включены.

Предлагаемое учебное пособие надо рассматривать как учебное руководство только по теоретической части основного втузовского курса математики. Для практических занятий учащиеся должны пользоваться рекомендованными сборниками задач.

Данное учебное пособие может быть использовано студентами дневных, вечерних и заочных втузов.

П. Романовский

ГЛАВА I

АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ НА ПЛОСКОСТИ

Раздел I

МЕТОД КООРДИНАТ

§ 1. Направленные отрезки

Одним из основных понятий аналитической геометрии является понятие *направленного отрезка*. Отрезок прямой, ограниченный точками A и B , называется *направленным*, если указано, какая из этих двух точек является его началом и какая — концом. Под *направлением* такого отрезка подразумевается направление от его начала к его концу. Будем обозначать направленный отрезок двумя буквами с общей стрелкой над ними. При этом сначала пишется буква, обозначающая начало отрезка, затем буква, обозначающая его конец.

Так, например, $\overrightarrow{AB}$ означает направленный отрезок с началом A и концом B , а $\overrightarrow{BA}$ означает направленный отрезок с началом B и концом A . Эти отрезки имеют одинаковую длину¹⁾ (обозначим ее AB или BA) и взаимно противоположные направления. Направленные отрезки используются во многих вопросах математики, физики и техники.

Возьмем произвольную прямую. На ней можно установить два взаимно противоположных направления. Выберем одно из них (любое) и назовем его *положительным* (а противоположное направление — *отрицательным*). Положительное направление принято на чертеже указывать стрелкой (рис. 1). Прямая с выбранным на ней направлением называется *осью*.

Рассмотрим направленные отрезки, лежащие на данной оси. Выбрав единицу масштаба, можно определить длины этих отрезков. *Величиной* любого направленного отрезка, лежащего на оси, называется его длина, взятая со знаком плюс, если направление отрезка совпадает с направлением оси, и со знаком минус, если эти направления противоположны. Величину направленного отрезка $\overrightarrow{AB}$ будем обозначать $\overline{AB}$.

В ряде случаев бывает удобно рассматривать также «нулевые» направленные отрезки, т. е. отрезки, у которых начало совпадает с концом. Длины и величины таких отрезков равны нулю, направления — произвольны.



Рис. 1.

¹⁾ Предполагается, что единица масштаба выбрана.

$$\overline{AB} = -\overline{BA}.$$
$$\overline{AB}=2, \quad \overline{AC}=-3, \quad \overline{BD}=-2+\sqrt{2}, \quad \overline{CD}=3+\sqrt{2}.$$
$$\overline{AC} = \overline{AB} + \overline{BC}. \quad (1)$$

A diagram showing a straight line with an arrow at the right end. Four points are marked on the line: C, A, D, and B, in that order from left to right. Below the line, there is a horizontal scale bar with vertical end caps, labeled with the letter 'e'.

$$\overline{AC} = \overline{AB} + \overline{BC}.$$
$$\overline{M_1 M_2} + \overline{M_2 M_3} + \dots + \overline{M_{n-1} M_n} = \overline{M_1 M_n}. \quad (2)$$
[illegible]

12

при измерении длин отрезков). Если на прямой выбрано направление, начальная точка O и единица масштаба, то говорят, что на этой прямой введена *декартова система координат*¹⁾. При этом сама прямая называется *координатной осью*, а точка O — *началом координат*.

Введение на прямой декартовой системы координат позволяет определить положение точек этой прямой с помощью действительных чисел.

Координатой любой точки M прямой в выбранной системе координат называется число x , равное величине направленного отрезка $\overrightarrow{OM}$ (рис. 4):

$$x = \overrightarrow{OM}.$$

Иначе, координатой точки M прямой называется число, равное по абсолютной величине расстоянию точки M от начала координат, положительное, если направление отрезка $\overrightarrow{OM}$ совпадает с направлением координатной оси, и отрицательное, если направление отрезка $\overrightarrow{OM}$ противоположно направлению этой оси.

Тот факт, что точка M имеет координату x , будем обозначать так: $M(x)$. Координата начала O равна нулю: $O(0)$.



Рис. 4.

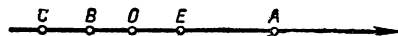


Рис. 5.

Если данная координатная ось расположена горизонтально, а положительное направление на ней выбрано слева направо (рис. 4), то любая точка M , лежащая справа от начала координат O , имеет положительную координату ($x > 0$), а любая точка M , лежащая слева от начала O , имеет отрицательную координату ($x < 0$).

На рис. 5 точки E, A, B, C имеют соответственно координаты 1, 3, -1 , -2 , т. е. $E(1)$, $A(3)$, $B(-1)$, $C(-2)$. Отрезок OE — масштабный отрезок.

Итак, если на прямой введена система координат, то каждой точке этой прямой соответствует определенное действительное число — координата этой точки.

Наоборот, каждое действительное число x является координатой одной и только одной точки данной прямой (в выбранной системе координат).

Для построения соответствующей числу x точки M достаточно построить направленный отрезок $\overrightarrow{OM}$, величина которого равна числу x .

Таким образом, при помощи декартовой системы координат на прямой можно установить взаимно однозначное соответствие между множеством всех точек прямой (координатной оси) и множеством всех действительных чисел: любой точке прямой соответствует определенное действительное число, а любому действительному числу — определенная точка на прямой.

В связи с этим в дальнейшем, имея в виду точки координатной оси, мы будем часто говорить о числах — их координатах. Например, говоря «дана точка», будем понимать, что дана ее координата; говоря, что надо «найти точку», будем искать ее координату; вместо того, чтобы на чертеже обозначать точку оси буквой, будем надписывать ее координату и т. д.

Считая заданной декартову систему координат на данной прямой, рассмотрим несколько задач.

Задача 1. Найти величину отрезка $\overline{M_1M_2}$, зная координату x_1 его начала и координату x_2 его конца.

¹⁾ Декарт Рене — французский математик и философ (1596—1650).

Решение. По условию $\overline{OM_1} = x_1$, $\overline{OM_2} = x_2$. Используя основное тождество (§ 1,1), имеем

$$\overline{M_1M_2} = \overline{M_1O} + \overline{OM_2},$$

т. е.

$$\overline{M_1M_2} = \overline{OM_2} - \overline{OM_1}.$$

Отсюда следует, что

$$\overline{M_1M_2} = x_2 - x_1. \quad (1)$$

Это равенство, как следует из доказательства, справедливо при любом расположении точек M_1 и M_2 .

Задача 2. Найти расстояние между точками M_1 и M_2 , зная их координаты x_1 и x_2 .

Решение. Расстояние между M_1 и M_2 (обозначим его d) равно длине отрезка $\overline{M_1M_2}$, следовательно, равно модулю величины этого отрезка. Таким образом,

$$d = |x_2 - x_1| \quad (2)$$

или, что то же, $d = |x_1 - x_2|$.

Например, если $M_1(-8)$, $M_2(2)$, то $d = M_1M_2 = |-8 - 2| = 10$.

Пусть M_1 и M_2 — две различные точки прямой, M_0 — произвольная (отличная от M_1 и M_2) точка той же прямой. *Отношением, в котором точка M_0 делит отрезок $\overline{M_1M_2}$, называется число λ , определяемое равенством*

$$\lambda = \frac{\overline{M_1M_0}}{\overline{M_0M_2}}. \quad (3)$$

Если точка M_0 совпадает с началом отрезка M_1 , то отношение считается равным нулю, если M_0 совпадает с концом M_2 — это отношение не определено.

Из определения следует, что число λ положительное, если отрезки $\overline{M_1M_0}$ и $\overline{M_0M_2}$ имеют одинаковое направление (точка M_0 лежит внутри отрезка $\overline{M_1M_2}$), и отрицательное, если отрезки $\overline{M_1M_0}$ и $\overline{M_0M_2}$ направлены противоположно (точка M_0 лежит вне отрезка $\overline{M_1M_2}$).¹⁾

Задача 3 (задача о делении отрезка в данном отношении). Зная точки $M_1(x_1)$ и $M_2(x_2)$ прямой и число λ ($\lambda \neq -1$), найти точку M , делящую отрезок $\overline{M_1M_2}$ в данном отношении λ .

Решение. Пусть x_0 — координата искомой точки M_0 .

По условию $\lambda = \frac{\overline{M_1M_0}}{\overline{M_0M_2}}$, или, так как $\overline{M_1M_0} = x_0 - x_1$, $\overline{M_0M_2} = x_2 - x_0$,

$$\lambda = \frac{x_0 - x_1}{x_2 - x_0}.$$

Решая это уравнение относительно x_0 , находим

$$x_0 = \frac{x_1 + \lambda x_2}{1 + \lambda}. \quad (4)$$

¹⁾ Если точка M_0 лежит внутри отрезка $\overline{M_1M_2}$, то

$$\lambda = \frac{M_1M_0}{M_0M_2},$$

а если точка M_0 лежит вне отрезка $\overline{M_1M_2}$, то

$$\lambda = -\frac{M_1M_0}{M_0M_2},$$

где M_1M_0 и M_0M_2 — длины соответствующих направленных отрезков.

Таким образом, поставленная задача для любого отрезка $\overline{M_1M_2}$ и любого числа λ ($\lambda \neq -1$) имеет, и притом единственное, решение. Соотношение (4) называют *формулой деления отрезка в данном отношении*.

Примечание. Если $\lambda = -1$, то задача не имеет решения, так как в этом случае

$$-1 = \frac{x_0 - x_1}{x_2 - x_0}, \text{ т. е. } x_1 = x_2,$$

что противоречит самой постановке задачи (отрезок $\overline{M_1M_2}$ — не нулевой!).

Пример. Разделить отрезок $\overline{M_1M_2}$, $M_1(1)$, $M_2(5)$, в отношении $\lambda = \frac{3}{2}$.

Решение.
$$x_0 = \frac{1 + \frac{3}{2} \cdot 5}{1 + \frac{3}{2}} = \frac{17}{5}.$$

Таким образом, точка $M_0\left(\frac{17}{5}\right)$ делит отрезок $\overline{M_1M_2}$ в отношении $\frac{3}{2}$.

Частным случаем рассмотренной выше задачи является задача о делении отрезка $\overline{M_1M_2}$ в отношении $\lambda = 1$. В этом случае

$$\frac{\overline{M_1M_0}}{\overline{M_0M_2}} = 1, \quad \overline{M_1M_0} = \overline{M_0M_2},$$

т. е. точка M_0 является серединой отрезка $\overline{M_1M_2}$.

Из формулы (4) получаем координату середины отрезка $\overline{M_1M_2}$:

$$x_0 = \frac{x_1 + x_2}{2}. \quad (5)$$

Так, например, серединой отрезка $\overline{M_1M_2}$, $M_1(3)$, $M_2(18)$, будет точка $M_0(x_0)$, где

$$x_0 = \frac{3 + 18}{2} = \frac{21}{2}.$$

§ 3. Прямоугольная декартова система координат на плоскости

В предыдущем параграфе мы рассматривали систему координат на прямой, позволяющую определять расположение точек на этой прямой с помощью чисел.

Покажем теперь, как с помощью пар чисел можно определить положение точек на плоскости.

Будем говорить, что на плоскости задана *декартова прямоугольная система координат*, если задан единичный (масштабный) отрезок, а также пара взаимно перпендикулярных осей и при этом условлено, которая из этих осей является первой и какая — второй.

Точку O пересечения этих осей будем называть *началом координат*, первую ось — *осью абсцисс*, или *осью* OX , вторую ось — *осью ординат*, или *осью* OY . Оси абсцисс и ординат называются *осями координат*. Обычно ось абсцисс располагают горизонтально, а ее положительным направлением считают направление слева направо, а ось OY — вертикально с положительным направлением снизу вверх ¹⁾. Вместо того, чтобы рисовать стрелки, указывающие направления осей, принято в положительных направлениях ставить соответственно буквы X и Y (рис. 6).

¹⁾ В дальнейшем, если не оговорено противное, расположение осей будет предполагаться именно таким, как указано на рис. 6.

Пусть M — произвольная точка плоскости. Опустим из точки M перпендикуляры на оси OX и OY . Точки пересечения этих перпендикуляров с осями координат обозначим соответственно через N и L (рис. 7).

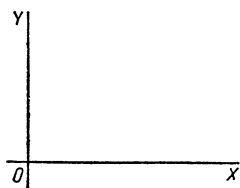


Рис. 6.

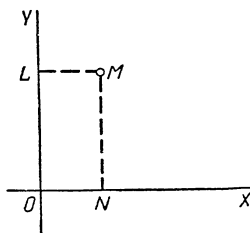


Рис. 7.

Абсциссой точки M будем называть величину отрезка $\overline{ON}$ оси OX , а ординатой — величину отрезка $\overline{OL}$ оси OY . Пару чисел x, y , где

$$x = \overline{ON}, \quad y = \overline{OL},$$

называют *координатами точки M* в выбранной системе координат. Тот факт, что точка M имеет координаты x и y , записывается так: $M(x, y)$. При этом сначала пишется абсцисса, затем ордината точки M .

Абсцисса x точки M равна нулю тогда и только тогда, когда точка M лежит на оси OY . Ордината y точки M равна нулю тогда и только тогда, когда точка M лежит на оси OX .

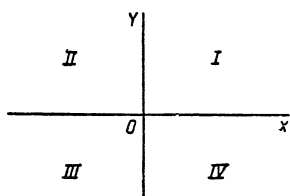


Рис. 8.

У начала координат O (и только у него) обе координаты равны нулю.

Ось абсцисс делит плоскость на две полуплоскости — *верхнюю* и *нижнюю*. У точек M , лежащих в верхней полуплоскости, ординаты положительны, а у точек нижней полуплоскости — отрицательны.

Ось ординат делит плоскость на две полуплоскости: *правую* и *левую*. У точек правой полуплоскости абсциссы положительны, у точек левой — отрицательны.

Оси координат делят плоскость на четыре части, называемые *четвертями*, *квадрантами* или *координатными углами*. Нумерация четвертей указана на рис. 8.

Возможны следующие комбинации знаков координат:

Четверть	I	II	III	IV
x	+	—	—	+
y	+	+	—	—

Таким образом, каждой точке M плоскости соответствует пара действительных чисел (x, y) — координат этой точки. Наоборот, каждой паре действительных чисел (x, y) соответствует, и притом одна, точка M плоскости, для которой эти числа будут ее координатами.

Построение точки M по ее координатам (x, y) осуществляется следующим образом. На оси OX строится отрезок $\overline{ON}$, величина которого равна x , на оси ординат строится отрезок $\overline{OL}$, величина которого равна y (при этом однозначно определяются точки N и L). В точках N и L восстанавливаются перпендикуляры к соответствующим осям. Точка пересечения

чения этих перпендикуляров имеет по определению координаты x и y , т. е. является искомой точкой M .

Таким образом, введение на плоскости декартовой прямоугольной системы координат позволяет установить взаимно однозначное соответствие между множеством всех точек плоскости и множеством пар¹⁾ действительных чисел. Это соответствие дает возможность сводить изучение множеств точек плоскости к изучению множеств пар действительных чисел, т. е. применять к изучению вопросов геометрии алгебраические методы. Это же соответствие позволяет геометрически излагать ряд вопросов алгебры, математического анализа и других дисциплин.

Условимся в дальнейшем, вместо того чтобы говорить «дана плоскость, на ней введена декартова прямоугольная система координат с осями OX и OY » говорить короче: «дана плоскость XOY » и считать, что «дана точка M », если даны ее координаты (x, y) (под отысканием точки понимается при этом отыскание ее координат).

§ 4. Простейшие задачи аналитической геометрии на плоскости

Задача 1. Дана точка $M(x, y)$. Найти точки, симметричные точке M относительно оси абсцисс, оси ординат и начала координат.

Решение. Обозначим искомые точки соответственно $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$ (рис. 9).

Так как точка M_1 симметрична точке M относительно оси абсцисс, то обозначая через N_1 и N проекции этих точек на ось абсцисс, а через L_1 и L — на ось ординат, получаем: $\overline{ON_1} = \overline{ON}$, $\overline{OL_1} = -\overline{OL}$, т. е. $x_1 = x$, $y_1 = -y$, $M_1(x, -y)$.

Рассуждая аналогичным образом, находим, что для точки M_2 , симметричной точке M относительно оси ординат, $x_2 = -x$, $y_2 = y$, т. е.

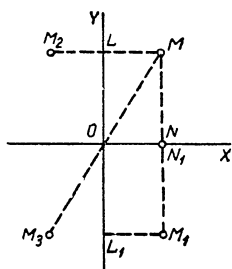


Рис. 9.

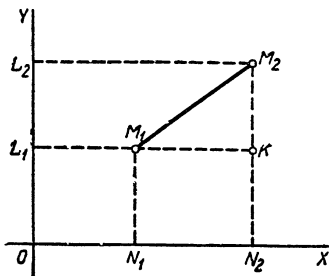


Рис. 10.

$M_2(-x, y)$, а для точки M_3 , симметричной точке M относительно начала координат, имеем $x_3 = -x$, $y_3 = -y$, т. е. $M_3(-x, -y)$.

Задача 2. Найти расстояние между двумя данными точками $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ плоскости.

Решение. Обозначим через N_1 и N_2 проекции точек M_1 и M_2 на ось OX , а через L_1 и L_2 — проекции M_1 и M_2 на ось OY . Обозначим затем через K точку пересечения прямых M_1L_1 и M_2N_2 (рис. 10).

По теореме Пифагора

$$M_1M_2 = \sqrt{M_1K^2 + KM_2^2}.$$

Так как

$$M_1K = |\overline{N_1N_2}| = |x_2 - x_1|,$$

$$KM_2 = |\overline{L_1L_2}| = |y_2 - y_1|,$$

¹⁾ Речь идет об упорядоченных парах действительных чисел, т. е. парах чисел, для которых известно, какое число является первым и какое — вторым.

то

$$M_1M_2 = \sqrt{|x_2 - x_1|^2 + |y_2 - y_1|^2}$$

или, обозначая длину отрезка $\overline{M_1M_2}$ через d :

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad (1)$$

Эта формула позволяет, зная координаты точек M_1 и M_2 , определять расстояние между ними.

Пример. Найти периметр $2p$ треугольника с вершинами $M_1(-6, -7)$, $M_2(-9, -3)$, $M_3(3, 2)$.

Решение. Так как

$$M_1M_2 = \sqrt{(-9+6)^2 + (-3+7)^2} = 5,$$

$$M_2M_3 = \sqrt{(3+9)^2 + (2+3)^2} = 13,$$

$$M_3M_1 = \sqrt{(-6-3)^2 + (-7-2)^2} = 9\sqrt{2},$$

то

$$2p = 5 + 13 + 9\sqrt{2} = 18 + 9\sqrt{2}.$$

Задача 3 (задача о делении отрезка в данном отношении). Пусть даны две различные точки плоскости M_1 и M_2 . Проведем через них прямую, выберем на ней направление и рассмотрим на ней точку M_0 , не совпадающую с точками M_1 и M_2 .

Отношением, в котором точка M_0 делит отрезок $\overline{M_1M_2}$, называется число λ , определяемое равенством

$$\lambda = \frac{\overline{M_1M_0}}{\overline{M_0M_2}}.$$

Если точка M_0 совпадает с точкой M_1 , то считается, что $\lambda = 0$. Если точка M_0 совпадает с точкой M_2 , то отношение не определено. Для точек M_0 , лежащих внутри

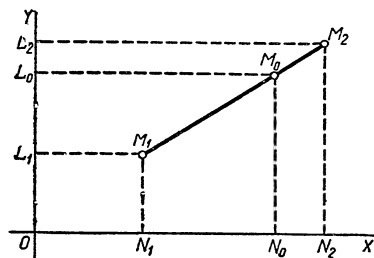


Рис. 11.

отрезка $\overline{M_1M_2}$, отношение λ положительно (про такие точки говорят, что они делят отрезок $\overline{M_1M_2}$ внутренним образом), для точек M_0 , лежащих на прямой вне отрезка $\overline{M_1M_2}$, отношение λ отрицательно (говорят, что эти точки делят отрезок $\overline{M_1M_2}$ в отношении λ внешним образом).

Так как точки M_1 и M_2 различны, то не существует точки M , для которой $\lambda = -1$ (см. стр. 15).

Рассмотрим теперь задачу о делении отрезка в данном отношении: зная точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ и число λ ($\lambda \neq -1$), определить на прямой M_1M_2 точку M_0 , делящую отрезок $\overline{M_1M_2}$ в отношении λ .

Решение. Обозначим через x_0 и y_0 координаты искомой точки M_0 , а через N_1 , N_2 и N_0 — проекции точек M_1 , M_2 и M_0 на ось OX (рис. 11).

Предположим, что прямая M_1M_2 не перпендикулярна оси OX . Прямые M_1N_1 , M_2N_2 и M_0N_0 параллельны, следовательно,

$$\frac{M_1M_0}{M_0M_2} = \frac{N_1N_0}{N_0N_2}.$$

Аналогичное равенство имеет место и для величин соответствующих направленных отрезков:

$$\frac{\overline{M_1M_0}}{\overline{M_0M_2}} = \frac{\overline{N_1N_0}}{\overline{N_0N_2}}.$$

Так как точки N_1 , N_2 и N_0 лежат на координатной оси — оси OX , то для нахождения координаты x_0 точки N_0 (она же — абсцисса точки M_0) применяем выведенную выше формулу (§ 2, 4):

$$x_0 = \frac{x_1 + \lambda x_2}{1 + \lambda}. \quad (2)$$

В случае перпендикулярности прямой M_1M_2 к оси OX формула (2) справедлива, ибо тогда $x_0 = x_1 = x_2$.

Аналогично получаем

$$y_0 = \frac{y_1 + \lambda y_2}{1 + \lambda}. \quad (3)$$

Формулы (2) и (3) называются *формулами деления направленного отрезка в данном отношении*.

В частности, если $\lambda = 1$, т. е. если

$$\frac{\overline{M_1M_0}}{\overline{M_0M_2}} = 1 \quad (\overline{M_1M_0} = \overline{M_0M_2}),$$

то точка M_0 есть середина отрезка $\overline{M_1M_2}$.

Ее координаты:

$$x_0 = \frac{x_1 + x_2}{2}, \quad (4)$$

$$y_0 = \frac{y_1 + y_2}{2}. \quad (5)$$

Пример 1. Даны две точки: $M_1(-8, 13)$ и $M_2(12, -17)$. Разделить отрезок $\overline{M_1M_2}$ в отношении $\lambda = \frac{3}{7}$.

Решение. Обозначая координаты искомой точки M_0 через x_0 , y_0 , получим:

$$x_0 = \frac{-8 + \frac{3}{7} \cdot 12}{1 + \frac{3}{7}} = -2; \quad y_0 = \frac{13 - \frac{3}{7} \cdot 17}{1 + \frac{3}{7}} = 4.$$

Пример 2. Найти центр тяжести однородного (по плотности) треугольника, заданного координатами своих вершин:

$$M_1(x_1, y_1), \quad M_2(x_2, y_2), \quad M_3(x_3, y_3).$$

Решение. Центр тяжести треугольника лежит в точке пересечения его медиан. Следовательно, для решения задачи достаточно найти точку пересечения двух любых медиан. Проведем медианы: M_3N , M_1L и обозначим искомую точку пересечения буквой K (рис. 12).

Найдем координаты точки $N(x_N, y_N)$ — середины отрезка $\overline{M_1M_2}$ по формулам (4), (5):

$$x_N = \frac{x_1 + x_2}{2}; \quad y_N = \frac{y_1 + y_2}{2}.$$

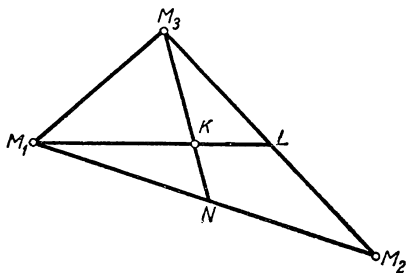


Рис. 12.

Теперь будем искать координаты точки $K(x_K, y_K)$ — точки, делящей отрезок $\overline{M_3N}$ в отношении $\lambda = \frac{2}{1}$ (на основании геометрических свойств медиан треугольника). Используя формулы деления отрезка в данном отношении (2), (3), получим:

$$x_K = \frac{x_3 + 2 \frac{x_1 + x_2}{2}}{1 + 2} = \frac{x_1 + x_2 + x_3}{3};$$

$$y_K = \frac{y_3 + 2 \frac{y_1 + y_2}{2}}{1 + 2} = \frac{y_1 + y_2 + y_3}{3}.$$

Итак,

$$K\left(\frac{x_1 + x_2 + x_3}{3}, \frac{y_1 + y_2 + y_3}{3}\right).$$

§ 5. Элементы теории проекций

Проекцией направленного отрезка $\overline{M_1M_2}$ на заданную ось l называется величина отрезка $\overline{N_1N_2}$, где N_1 и N_2 — проекции точек M_1 и M_2 на ось l (рис. 13)¹⁾. По определению

$$\text{пр}_l \overline{M_1M_2} = \overline{N_1N_2}.$$

Таким образом, проекция направленного отрезка на ось есть число (оно может быть положительным, отрицательным и нулем).

Рассмотрим на плоскости направленные отрезки

$$\overline{M_1M_2}, \overline{M_2M_3}, \dots, \overline{M_{n-1}M_n}$$

и составленную из них ломаную, обозначим ее $\overline{M_1M_2 \dots M_n}$ (рис. 14).

Эту ломаную будем называть *направленной*, а точки M_1 и M_n соответственно ее *началом* и *концом*.

Проекцией направленной ломаной линии $\overline{M_1M_2 \dots M_n}$ на заданную ось l называется проекция замыкающего эту ломаную отрезка $\overline{M_1M_n}$ на

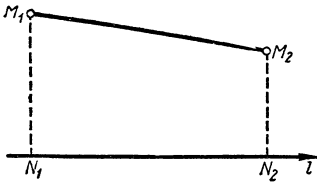


Рис. 13.

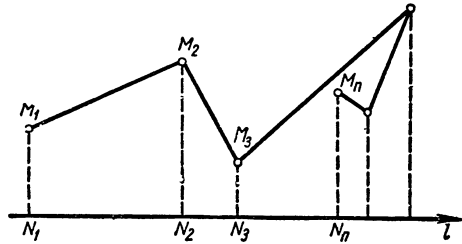


Рис. 14.

ось l . Из этого определения следует, что проекция направленной ломаной на ось зависит только от положения начальной и конечной точки ломаной. В частности, проекция замкнутой направленной ломаной всегда равна нулю.

На основании тождества (§ 1, 2) непосредственно заключаем, что проекция любой направленной ломаной равна алгебраической сумме проекций ее звеньев.

В самом деле, если N_k обозначает проекцию точки M_k на ось l , $k = 1, 2, \dots, n$, то

$$\text{пр}_l \overline{M_1M_2 \dots M_n} = \overline{N_1N_n}, \quad (1)$$

$\text{пр}_l \overline{M_1M_2} + \text{пр}_l \overline{M_2M_3} + \dots + \text{пр}_l \overline{M_{n-1}M_n} = \overline{N_1N_2} + \overline{N_2N_3} + \dots + \overline{N_{n-1}N_n}$,
Так как правые части этих равенств равны (см. § 1, 2), то равны и левые:

$$\text{пр}_l \overline{M_1M_2 \dots M_n} = \text{пр}_l \overline{M_1M_2} + \text{пр}_l \overline{M_2M_3} + \dots + \text{пр}_l \overline{M_{n-1}M_n}.$$

Угол между осями. Пусть на плоскости заданы две оси l_1 и l_2 , пересекающиеся в некоторой точке S . Под *углом между* этими осями, заданными в определенном порядке, понимается угол, на который надо повернуть вокруг точки S первую ось (l_1) до совпадения ее положительного направления с положительным направлением второй оси (l_2). Этот угол будем обозначать (l_1, l_2) и измерять в градусах или в радианах. Угол (l_1, l_2) будем считать положительным, если поворот оси l_1 до совпадения ее направления с направлением оси l_2 производится против часовой стрелки, и *отри-*

¹⁾ Напомним, что проекция точки на ось есть основание перпендикуляра, опущенного из этой точки на ось.

цательным — в противном случае. Величина угла между осями определяется неоднозначно, так как, повернув ось l_1 на 360° по или против часовой стрелки, мы получим прежнее ее положение. Если φ — какой-нибудь угол между осями l_1 и l_2 , измеренный в радианах, то число $\varphi + 2k\pi$, где k — любое целое число, также равно углу между этими осями.

Если оси l_1 и l_2 параллельны и одинаково направлены, то угол между осями считаем равным нулю (или $2k\pi$), а если оси l_1 и l_2 параллельны и противоположно направлены, угол между ними считаем равным π (или $\pi + 2k\pi$)¹⁾.

Под углом между осью l и направленным отрезком $\overline{M_1M_2}$ будем понимать угол между осью l и осью, проходящей через точки M_1 и M_2 в направлении от M_1 к M_2 (рис. 15). Угол между осью l и направлен-

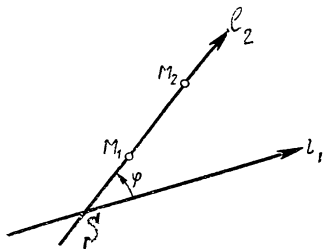


Рис. 15.

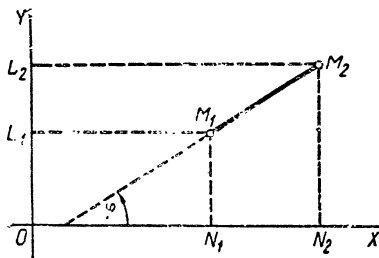


Рис. 16.

ным отрезком $\overline{M_1M_2}$ называют также *углом наклона* этого отрезка к оси l ²⁾.

Проекция направленного отрезка на оси координат. Если направленный отрезок $\overline{M_1M_2}$ имеет угол наклона φ к некоторой оси l , то

$$\text{пр}_l \overline{M_1M_2} = M_1M_2 \cos \varphi^3). \quad (2)$$

Пусть на плоскости XOY даны точки $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ и известен угол φ наклона отрезка $\overline{M_1M_2}$ к оси OX (рис. 16).

Проектируя отрезок $\overline{M_1M_2}$ на ось абсцисс, получим:

$$\text{пр}_{OX} \overline{M_1M_2} = \overline{N_1N_2} = M_1M_2 \cos \varphi. \quad (3)$$

Проектируя отрезок $\overline{M_1M_2}$ на ось ординат, аналогично получим:

$$\text{пр}_{OY} \overline{M_1M_2} = \overline{L_1L_2} = M_1M_2 \cos \left(\varphi - \frac{\pi}{2} \right) = M_1M_2 \sin \varphi. \quad (4)$$

Так как абсциссы точек M_1 и N_1 , M_2 и N_2 соответственно равны, то на основании равенства (§ 2,1) можно записать, что

$$\overline{N_1N_2} = x_2 - x_1, \quad (5)$$

а так как ординаты точек M_1 и L_1 , M_2 и L_2 соответственно равны, то аналогично можно записать, что

$$\overline{L_1L_2} = y_2 - y_1. \quad (6)$$

¹⁾ Заметим еще, что если угол между осями l_1 и l_2 равен φ , то после поворота оси l_1 на угол α в направлении против часовой стрелки упомянутый угол «уменьшится» на α , так как станет равным $\varphi - \alpha$, а после поворота оси l_2 на угол α в направлении против часовой стрелки угол между осями l_1 и l_2 «увеличится» на α , т. е. станет равным $\varphi + \alpha$.

²⁾ Аналогично определяется угол наклона луча (полуоси) к оси l .

³⁾ Эта формула допускает очевидное обобщение. Пусть l_1 и l_2 две оси, φ угол между l_1 и l_2 , M_1 и M_2 любые точки на l_2 , тогда $\text{пр}_{l_1} \overline{M_1M_2} = \overline{M_1M_2} \cos \varphi$.

Обозначая длину отрезка $\overline{M_1M_2}$ через d и используя равенства (5) и (6), запишем равенства (3) и (4) в следующем виде:

$$x_2 - x_1 = d \cos \varphi, \quad (7)$$

$$y_2 - y_1 = d \sin \varphi. \quad (8)$$

Формулы (7) и (8) выражают проекции произвольного отрезка на координатные оси через его длину и угол наклона отрезка к оси абсцисс¹⁾.

В случае, если $x_2 - x_1 \neq 0$ (отрезок $\overline{M_1M_2}$ не параллелен оси ординат), из соотношений (7), (8) вытекает равенство

$$\operatorname{tg} \varphi = \frac{y_2 - y_1}{x_2 - x_1}, \quad (9)$$

которое может быть использовано для определения угла наклона отрезка $\overline{M_1M_2}$ к оси OX по известным координатам начала и конца этого отрезка.

§ 6. Полярная система координат

В предшествующих параграфах для определения положения точки на плоскости использовалась декартова прямоугольная система координат. Существуют и другие системы координат, позволяющие определять положение точки на плоскости с помощью пар действительных чисел. К таким системам координат относится, например, рассмотренная ниже полярная система координат.

Возьмем на плоскости точку O и проходящую через нее ось OP . Будем называть точку O *полюсом*, а полуось (луч), выходящую из точки O в положительном для оси OP направлении, — *полярной осью*. Задание полюса O , полярной оси OP и единичного (масштабного) отрезка OE определяет на плоскости *полярную систему координат*.

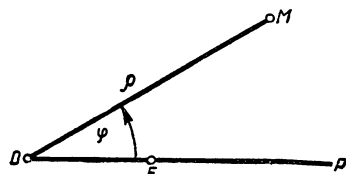


Рис. 17.

Полярным радиусом любой точки M плоскости называется ее расстояние от полюса O , т. е. длина отрезка $\overline{OM}$. *Полярным углом* φ точки M ²⁾ называется угол наклона направленного отрезка $\overline{OM}$ к полярной оси OP (рис. 17). Угол φ определяется с учетом знака и с точностью до слагаемого вида $2k\pi$, где k — целое число (см. § 5). Обычно в качестве полярных углов точек плоскости берут так называемые *главные их значения*, т. е. значения, определяемые равенствами: $0 \leq \varphi < 2\pi$ ³⁾.

Числа ρ и φ , т. е. полярный радиус и полярный угол точки M , называются ее *полярными координатами*. Точка с полярными координатами ρ и φ обозначается так: $M(\rho, \varphi)$ или (ρ, φ) ⁴⁾.

Таким образом, любой точке M плоскости (кроме полюса) соответствует определенная (если берется главное значение полярного угла) пара действительных чисел — ее полярных координат. Для точки O (полюса) по определению первая полярная координата ρ равна нулю, а вторая — угол φ — не имеет определенного значения.

¹⁾ Эти формулы допускают следующие обобщения: если на какой-нибудь оси, наклоненной к оси OX под углом φ , взяты любые точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$, то

$$\begin{cases} x_2 - x_1 = \overline{M_1M_2} \cos \varphi, \\ y_2 - y_1 = \overline{M_1M_2} \sin \varphi. \end{cases}$$

²⁾ Точку M считаем отличной от полюса O .

³⁾ Иногда главными считают значения φ в пределах: $-\pi < \varphi \leq \pi$.

⁴⁾ На первом месте пишется полярный радиус, на втором — угол.

Задание любой пары действительных чисел (ρ, φ) , $\rho \geq 0$, позволяет построить на плоскости одну (и только одну) точку M , для которой эти числа являются ее полярными координатами. Если $\rho = 0$, то точка M совпадает с полюсом O . Если же $\rho > 0$, то построение точки M сводится к построению отрезка $\overline{OM}$, угол наклона которого к полярной оси равен φ ¹⁾, а длина равна ρ .

Пример. Построить точки $M_1\left(2, \frac{3\pi}{4}\right)$, $M_2\left(2, -\frac{\pi}{6}\right)$, $M_3\left(\frac{3}{2}, \frac{17\pi}{4}\right)$ в заданной полярной системе координат.

Решение. Проведем через полюс O луч (полуось) под углом $\frac{3\pi}{4}$ к полярной оси (иначе, повернем полярную ось на угол $\frac{3\pi}{4}$ вокруг точки O против часовой стрелки), затем отложим на полученном луче от полюса O (в положительном направлении) отрезок длины, равной 2. Его конец — точка M_1 (рис. 18).

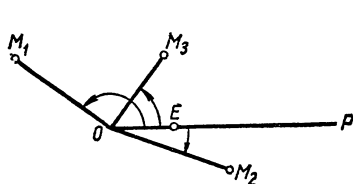


Рис. 18.

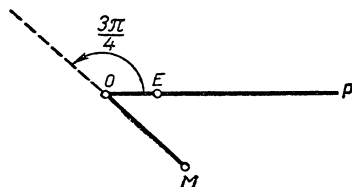


Рис. 19.

Для построения точки M_2 надо провести луч под углом $-\frac{\pi}{6}$ к полярной оси и отложить на нем от точки O две единицы масштаба.

Для построения точки M_3 нужно провести луч, составляющий с полярной осью угол $\frac{17\pi}{4}$ или, что то же, $\frac{\pi}{4}$ (так как $\frac{17\pi}{4} = \frac{\pi}{4} + 4\pi$), а затем отложить на нем из точки O отрезок, длина которого $\frac{3}{2}$.

Примечание. В некоторых случаях ограничение $\rho \geq 0$, накладываемое на полярный радиус точки, оказывается по тем или иным причинам неудобным. Иногда условливаются допускать значения $\rho < 0$, понимая при этом под точкой (ρ, φ) точку $(|\rho|, \varphi + \pi)$.

Для построения точки (ρ, φ) , где $\rho < 0$, нужно построить луч, имеющий к полярной оси угол наклона φ , а затем отложить от полюса на продолжении луча отрезок, длина которого равна $|\rho|$.

На рис. 19 построены точки E и M с полярными координатами соответственно $(1, 0)$ и $\left(-2, \frac{3\pi}{4}\right)$.

Введение обобщенной полярной системы координат позволяет каждой паре действительных чисел поставить в соответствие одну определенную точку плоскости.

§ 7. Связь между прямоугольными декартовыми и полярными координатами

В некоторых случаях приходится пользоваться одновременно декартовыми прямоугольными и полярными координатами точек. В связи с этим представляют интерес формулы, позволяющие по декартовым координатам точки находить полярные координаты и наоборот.

¹⁾ Напомним, что если $\varphi > 0$, то направление $\overline{OM}$ получается поворотом полярной оси вокруг точки O против часовой стрелки, а если $\varphi < 0$, то по часовой стрелке.

Рассмотрим на плоскости прямоугольную декартову систему координат XOY и полярную систему координат, у которой полюс совпадает с началом координат O , а полярная ось совпадает с осью абсцисс¹⁾.

Пусть M — произвольная точка плоскости (отличная от полюса), x, y — ее декартовы координаты, а ρ, φ — полярные (рис. 20).

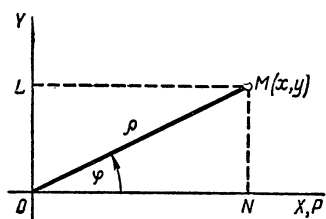


Рис. 20.

Так как

$$\cos \varphi = \frac{x}{\rho}, \quad \sin \varphi = \frac{y}{\rho},$$

то

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi. \quad (1)$$

Формулы (1) выражают прямоугольные декартовы координаты точки M через полярные координаты.

Далее, при любом положении точки M

$$\rho^2 = |x|^2 + |y|^2, \quad \text{т. е.} \quad \rho^2 = x^2 + y^2,$$

следовательно,

$$\rho = \sqrt{x^2 + y^2}, \quad \cos \varphi = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \varphi = \frac{y}{\sqrt{x^2 + y^2}}. \quad (2)$$

Формулы (2) позволяют определить полярные координаты точки M по ее декартовым координатам.

Если точка M не лежит на оси OY , то из формул (2) следует также соотношение

$$\operatorname{tg} \varphi = \frac{y}{x}.$$

§ 8. Преобразование координат

Меняя направления осей, положения начала координат и полюса, изменяя единицы масштаба, можно получать на плоскости различные декартовы и полярные системы координат. Фиксируем произвольную точку M плоскости. В различных системах координат эта точка будет иметь, вообще говоря, различные координаты. Возникает вопрос: как связаны между собой координаты точки M в различных системах координат? Или иначе, каким образом, зная координаты точки M в одной системе координат, найти ее координаты в другой системе? Ответ на этот вопрос дают так называемые *формулы преобразования координат*. С формулами, связывающими декартовы и полярные координаты точки, мы познакомились в конце предыдущего параграфа. Ниже рассматриваются некоторые другие формулы преобразования координат.

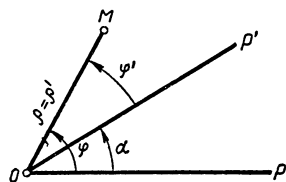


Рис. 21.

1. Поворот полярной оси. Рассмотрим две полярные системы координат с общим полюсом O , одинаковыми масштабными отрезками, но разными направлениями полярных осей OP и OP' . Обозначим через α угол, составляемый осью OP с осью OP' (рис. 21). Будем говорить, что полярная система координат с осью OP' (новая система) получена из полярной системы координат с осью OP (старой системы) поворотом оси вокруг полюса на угол α . Пусть M — произвольная точка плоскости, ρ, φ и ρ', φ' — ее старые и новые полярные координаты.

¹⁾ Единицы масштаба по осям декартовой и полярной систем координат считаются одинаковыми.

Очевидно, $\rho = \rho'$, $\varphi = \varphi' + \alpha$
или, что то же,

$$\rho' = \rho, \quad \varphi' = \varphi - \alpha. \quad (1)$$

Формулы (1) устанавливают связь между старыми и новыми полярными координатами точки M .

2. Параллельный перенос осей прямоугольной декартовой системы координат. Рассмотрим две прямоугольные декартовы системы координат XOY и $X'O'Y'$, у которых направления соответствующих осей одинаковы, масштабные отрезки равны, а начала координат O и O' различны (рис. 22).

В этом случае система координат $X'O'Y'$ (назовем ее новой) получается из системы XOY (старой) *параллельным переносом* (параллельным сдвигом) *осей координат*¹⁾.

Положение новой системы координат относительно старой в данном случае полностью определяется заданием координат нового начала O' в старой системе координат: $O'(a, b)$.

Пусть M — произвольная точка плоскости, x, y и x', y' — ее координаты в системах XOY и $X'O'Y'$. Спроектировав точку M на старые и новые оси координат, получим точки N, L, N', L' . При этом

$$x = \overline{ON}, \quad y = \overline{OL}, \quad x' = \overline{O'N'}, \quad y' = \overline{O'L'}.$$

Обозначая через A и B проекции точки O' на оси OX и OY и учитывая основное тождество, будем иметь:

$$\begin{aligned} \overline{ON} &= \overline{OA} + \overline{AN}, \\ \overline{OL} &= \overline{OB} + \overline{BL} \end{aligned}$$

или, так как $\overline{AN} = \overline{O'N'}$, $\overline{BL} = \overline{O'L'}$:

$$\begin{aligned} \overline{ON} &= \overline{OA} + \overline{O'N'}, \\ \overline{OL} &= \overline{OB} + \overline{O'L'}. \end{aligned}$$

Учитывая, что $\overline{OA} = a$, $\overline{OB} = b$, получаем формулы:

$$\begin{aligned} x &= x' + a, \\ y &= y' + b, \end{aligned} \quad (2)$$

выражающие старые координаты точки M через ее новые координаты. При этом новые координаты точки выражаются через старые по формулам:

$$\begin{aligned} x' &= x - a, \\ y' &= y - b. \end{aligned} \quad (3)$$

Равенства (2) и (3) представляют собой формулы преобразования координат произвольной точки M плоскости при параллельном переносе осей координат (коротко, формулы параллельного переноса осей координат).

3. Поворот осей декартовой системы координат. Даны две декартовы прямоугольные системы координат XOY и $X'O'Y'$ с общим началом, одинаковыми масштабными отрезками и разными направлениями осей (ось OX составляет с осью OX' угол α). Будем говорить, что новая система координат $X'O'Y'$ получена из старой системы XOY поворотом

¹⁾ Система XOY при этом также получается из системы $X'O'Y'$ путем параллельного переноса осей координат.

осей координат вокруг начала на угол α . Пусть M — произвольная точка плоскости, x, y и x', y' — ее координаты соответственно в старой и новой системах координат. Спроектируем точку M на оси координат, при этом $\overline{ON} = x, \overline{OL} = y, \overline{ON'} = x', \overline{OL'} = y'$ (рис. 23).

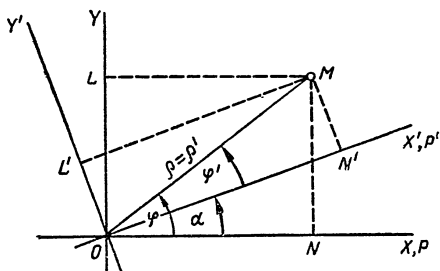


Рис. 23.

получена из полярной системы с осью OP поворотом оси OP вокруг точки O на угол α , то

$$\rho = \rho', \quad \varphi = \varphi' + \alpha.$$

Применяя формулы перехода от полярных координат к декартовым, будем иметь:

$$\begin{aligned} x &= \rho \cos \varphi, & y &= \rho \sin \varphi, \\ x' &= \rho \cos \varphi', & y' &= \rho \sin \varphi'. \end{aligned}$$

Так как $\varphi = \varphi' + \alpha$, то

$$\begin{aligned} x &= \rho \cos(\varphi' + \alpha) = \rho (\cos \varphi' \cos \alpha - \sin \varphi' \sin \alpha) = x' \cos \alpha - y' \sin \alpha, \\ y &= \rho \sin(\varphi' + \alpha) = \rho (\sin \varphi' \cos \alpha + \cos \varphi' \sin \alpha) = y' \cos \alpha + x' \sin \alpha. \end{aligned}$$

Окончательно получаем следующие формулы:

$$\begin{aligned} x &= x' \cos \alpha - y' \sin \alpha, \\ y &= x' \sin \alpha + y' \cos \alpha, \end{aligned} \quad (4)$$

выражающие старые координаты точки M через новые.

Так как система координат XOY , в свою очередь, получается из системы $X'OY'$ поворотом осей вокруг начала на угол $-\alpha$, то из формул (4), меняя соответственно обозначения координат, получим равенства

$$\begin{aligned} x' &= x \cos \alpha + y \sin \alpha, \\ y' &= -x \sin \alpha + y \cos \alpha, \end{aligned} \quad (5)$$

выражающие новые координаты точки M через старые.

Равенства (4) и (5) являются формулами преобразования координат произвольной точки M при повороте осей координат (для краткости их называют *формулами поворота осей координат*).

4. Изменение начала и поворот осей декартовой прямоугольной системы координат. Рассмотрим две прямоугольные декартовы системы координат XOY и $X'O'Y'$ с различными началами координат и различным направлением осей координат (но с одинаковыми масштабными отрезками!). Пусть новое начало O' в старой системе координат XOY имеет координаты a, b , а ось OX составляет с осью $O'X'$ угол α (рис. 24).

Введем вспомогательную систему координат $X''O'Y''$, у которой начало O' , а оси $O'X''$ и $O'Y''$ имеют одинаковые направления с осями OX и OY .

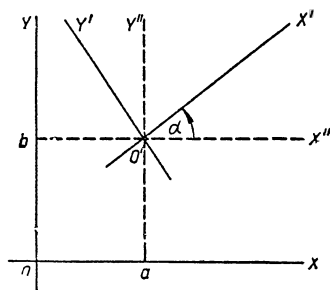


Рис. 24.

Переход от системы координат XOY к системе $X'O'Y'$ можно совершить последовательно, переходя от данной системы к системе $X''O'Y''$ (параллельный перенос осей координат), а затем от системы $X''O'Y''$ к системе $X'O'Y'$ (поворот осей координат).

Обозначая через $x, y; x', y'$ и x'', y'' координаты произвольно взятой точки M в соответствующих системах координат и используя выведенные выше формулы, получим:

$$\begin{aligned}x &= x'' + a, & y &= y'' + b, \\x'' &= x' \cos \alpha - y' \sin \alpha, & y'' &= x' \sin \alpha + y' \cos \alpha.\end{aligned}$$

Отсюда следует, что

$$\begin{aligned}x &= x' \cos \alpha - y' \sin \alpha + a, \\y &= x' \sin \alpha + y' \cos \alpha + b.\end{aligned}\tag{6}$$

Формулы (6) дают выражения старых координат точки M через новые.

Выражения новых координат точки M через старые

$$\begin{aligned}x' &= (x - a) \cos \alpha + (y - b) \sin \alpha, \\y' &= -(x - a) \sin \alpha + (y - b) \cos \alpha\end{aligned}\tag{7}$$

непосредственно следуют из соотношений (6).

Примечание. Формулы (6) и (7), как частный случай, включают в себя формулы (2), (3), (4), (5).

Пример. Зная координаты точки M (1, 3) в декартовой системе координат XOY , найти ее координаты в новой системе координат $X'O'Y'$, полученной из данной переносом начала O в точку O' (2, -1) и поворотом осей на угол $+\frac{\pi}{4}$.

Решение. Полагая в формулах (7) $a=2, b=-1, \alpha=\frac{\pi}{4}$, получим формулы преобразования координат, соответствующие данным условиям:

$$\begin{aligned}x' &= (x - 2) \frac{\sqrt{2}}{2} + (y + 1) \frac{\sqrt{2}}{2}, \\y' &= -(x - 2) \frac{\sqrt{2}}{2} + (y + 1) \frac{\sqrt{2}}{2}.\end{aligned}$$

Подставляя в них $x=1, y=3$; найдем новые координаты точки M :

$$\begin{aligned}x' &= (1 - 2) \frac{\sqrt{2}}{2} + (3 + 1) \frac{\sqrt{2}}{2} = \frac{3\sqrt{2}}{2}, \\y' &= -(1 - 2) \frac{\sqrt{2}}{2} + (3 + 1) \frac{\sqrt{2}}{2} = \frac{5\sqrt{2}}{2}.\end{aligned}$$

Раздел 2

ЛИНИИ И ИХ УРАВНЕНИЯ

§ 9. Линии и их уравнения

1. Об уравнениях, связывающих две переменные величины. Обозначим через x и y переменные величины, каждая из которых может принимать различные действительные значения.

Рассмотрим уравнение:

$$F(x, y) = 0,\tag{1}$$

где $F(x, y)$ — выражение, содержащее x и y . Например, $x + 2y + xy^2 = 0$ [здесь $F(x, y) = x + 2y + xy^2$], $(x + y)(x - y) = 0$ [здесь $F(x, y) = (x + y)(x - y)$] и т. д.

Будем говорить, что числа x_0, y_0 удовлетворяют уравнению (1), если подставив их вместо переменных x и y в выражение $F(x, y)$, мы получим тождество

$$F(x_0, y_0) = 0.$$

Числа x_0, y_0 не удовлетворяют уравнению (1), если подставляя их вместо переменных x и y в левую часть этого уравнения, мы получим:

$$F(x_0, y_0) \neq 0.$$

Например, числа $x=1, y=3$ удовлетворяют уравнению $3x - y = 0$, так как $3 \cdot 1 - 3 = 0$, а числа $x=5, y=4$ этому уравнению не удовлетворяют, так как $3 \cdot 5 - 4 \neq 0$.

Уравнению

$$F(x, y) = 0$$

может удовлетворять одна пара действительных чисел, несколько и даже бесчисленное множество таких пар.

Например, уравнению $(x-2)^2 + (y-4)^2 = 0$ удовлетворяет единственная пара чисел: $x=2, y=4$, уравнению $[(x-1)^2 + (y+1)^2] \times [x^2 + y^2] = 0$ удовлетворяют две пары чисел: $x=1, y=-1$ и $x=0, y=0$. Уравнению $x + 2y - 1 = 0$ удовлетворяют пары чисел $x=1, y=0$; $x=3, y=-1$ и вообще, всякая пара чисел $x=x_0, y=\frac{1-x_0}{2}$, где x_0 — любое действительное число.

Существуют уравнения, которым не удовлетворяет ни одна пара действительных чисел. Таковыми будут, например, уравнения $x^4 + y^4 + 1 = 0$, $\sin(x+y) - 3 = 0$.

2. Уравнение плоской линии. В аналитической геометрии линии рассматриваются как геометрические места точек, их составляющих. Например, окружность определяется как геометрическое место точек плоскости, равноотстоящих от некоторой фиксированной точки плоскости (центра окружности), биссектрису плоского угла можно рассматривать как геометрическое место точек, равноотстоящих от сторон этого угла и т. д.

В настоящем разделе будут изучаться только *плоские линии* (или, как их еще называют, *плоские кривые*), т. е. линии, лежащие в некоторой плоскости. Введение на плоскости системы координат (декартовой или полярной) позволяет положение точек плоскости однозначно определять заданием пар действительных чисел — их координат, а положение линий на плоскости определять с помощью равенств, связывающих пары чисел — координаты точек этих линий, т. е. с помощью уравнений.

Пусть на данной плоскости выбрана декартова прямоугольная система координат $ХОУ$.

Уравнение

$$F(x, y) = 0, \tag{1}$$

связывающее две переменные величины x и y , называется *уравнением линии* L (в выбранной на плоскости системе координат), если координаты любой точки линии L удовлетворяют этому уравнению, а координаты точек, не принадлежащих линии L , этому уравнению не удовлетворяют.

Таким образом, уравнение линии L есть соотношение, связывающее координаты точек данной линии (и только этих точек). Это соотношение представляет собой аналитическую запись, т. е. запись с помощью «формулы», того свойства, которое выделяет среди всех точек плоскости

точки данной линии. Так как линия рассматривается как геометрическое место точек, то можно сказать, что уравнение линии представляет собой запись свойства, определяющего данное геометрическое место точек.

Так, например, окружность радиуса R с центром в точке $O'(a, b)$ определяется как геометрическое место точек плоскости, находящихся от точки O' на расстоянии R (рис. 25).

Уравнение этой окружности имеет вид:

$$(x-a)^2 + (y-b)^2 - R^2 = 0 \quad (2)$$

или, что то же,

$$(x-a)^2 + (y-b)^2 = R^2. \quad (2')$$

Действительно, пусть $M(x, y)$ — любая точка окружности. Тогда ее расстояние от точки $O'(a, b)$ равно числу R :

$$\sqrt{(x-a)^2 + (y-b)^2} = R.$$

При этом $(x-a)^2 + (y-b)^2 = R^2$, а значит, координаты точки $M(x, y)$ удовлетворяют уравнению (2). Пусть теперь $M(x, y)$ — любая

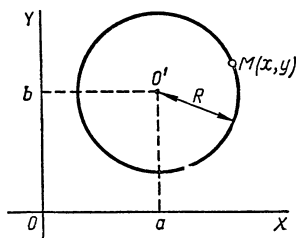


Рис. 25.

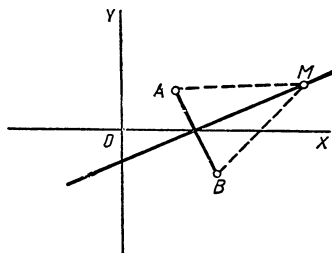


Рис. 26.

точка, не принадлежащая данной окружности. Тогда ее расстояние от точки O' либо больше, либо меньше R :

$$\sqrt{(x-a)^2 + (y-b)^2} > R \quad \text{или} \quad \sqrt{(x-a)^2 + (y-b)^2} < R.$$

Для такой точки

$$(x-a)^2 + (y-b)^2 > R^2 \quad \text{или} \quad (x-a)^2 + (y-b)^2 < R^2,$$

т. е. ее координаты не удовлетворяют уравнению (2). Следовательно, по определению, уравнение (2) есть уравнение данной окружности.

Рассмотрим пример на составление уравнения линии. Пусть требуется найти уравнение прямой, делящей пополам отрезок между точками $A(1, 1)$ и $B(2, -1)$ и перпендикулярной к этому отрезку (рис. 26).

Данная прямая представляет собой геометрическое место точек плоскости, равноудаленных от точек A и B . Пусть $M(x, y)$ — произвольная точка этой прямой. Тогда по условию

$$MA = MB,$$

т. е.

$$\sqrt{(x-1)^2 + (y-1)^2} = \sqrt{(x-2)^2 + (y+1)^2}.$$

Упрощая полученное соотношение, приходим к уравнению

$$-2x + 4y + 3 = 0,$$

которому удовлетворяют координаты x, y точки M данной прямой.

Если точка $M(x, y)$ не лежит на данной прямой, то либо $MA > MB$, либо $MA < MB$. В обоих случаях координаты точки M уравнению $-2x + 4y + 3 = 0$ удовлетворять не будут.

Аналогичным образом вводится понятие уравнения линии в выбранной на плоскости полярной системе координат.

Уравнение $F(\rho, \varphi) = 0$ называется *уравнением* данной линии в некоторой полярной системе координат, если координаты любой точки, лежащей на этой линии (и только они) удовлетворяют этому уравнению.

Пусть, например, требуется составить в данной полярной системе координат уравнение окружности, проходящей через полюс, с центром в точке $O'(R, 0)$ полярной оси (рис. 27).

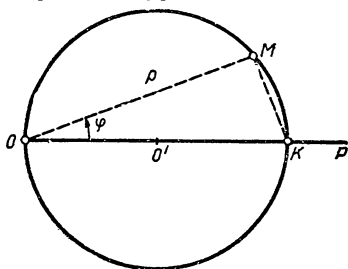


Рис. 27.

Пусть M — любая точка окружности, ρ и φ — ее полярные координаты. Рассмотрим треугольник OMK , где K — отличная от полюса точка пересечения окружности с полярной осью. Так как угол OMK прямой, то $\rho = 2R \cos \varphi$.

Уравнение $\rho = 2R \cos \varphi$ или, что то же, $\rho - 2R \cos \varphi = 0$ является уравнением окружности в выбранной полярной системе координат.

Действительно, координаты любой точки $M(\rho, \varphi)$ окружности удовлетворяют этому уравнению.

Если же точка $M(\rho, \varphi)$ лежит вне данной окружности, то угол OMK не будет прямым, т. е. соотношение $\rho = 2R \cos \varphi$ не будет иметь места.

Примечание. Если центр окружности радиуса R находится в полюсе O , то уравнение окружности, очевидно, имеет вид: $\rho = R$.

Исследуя уравнения линий в выбранных системах координат, можно делать различные заключения о расположении и свойствах этих линий. Таким образом, введение на плоскости системы координат позволяет изучение геометрических свойств линий заменить исследованием соответствующих им уравнений.

Так, для того чтобы ответить на вопрос, лежит ли точка M_0 на данной линии, достаточно проверить, удовлетворяют ли координаты точки M_0 уравнению этой линии в выбранной системе координат.

Пример. Точка $M_0(4, 1)$ лежит на окружности радиуса $R=3$ с центром в точке $O'(1, 1)$, так как ее координаты 4, 1 удовлетворяют уравнению этой окружности: $(x-1)^2 + (y-1)^2 = 9$.

Действительно, $(4-1)^2 + (1-1)^2 = 9$. Точка $M_1(-1, 3)$ на данной окружности не лежит, так как

$$(-1-1)^2 + (3-1)^2 \neq 9.$$

Задача о нахождении точек пересечения двух линий, имеющих в выбранной системе координат уравнения $F(x, y) = 0$ и $\Phi(x, y) = 0$, сводится к отысканию точек, координаты которых удовлетворяют уравнениям обеих кривых, т. е. сводится к решению системы двух уравнений с двумя неизвестными:

$$\begin{aligned} F(x, y) &= 0, \\ \Phi(x, y) &= 0. \end{aligned} \tag{3}$$

Примечание. Решением системы уравнений (3) в алгебре называют всякую пару действительных чисел $x=x_0, y=y_0$, удовлетворяющую каждому из уравнений.

Система (3) может иметь одно решение $x=x_0, y=y_0$ [данные кривые пересекаются в одной точке $M_0(x_0, y_0)$], может иметь несколько и даже бесчисленное множество решений (кривые будут пересекаться соответственно в нескольких точках или в бесчисленном множестве точек. Может также случиться, что система (3) не имеет решения. В этом случае данные кривые не имеют общих точек, т. е. не пересекаются.

Пример. Требуется найти точки пересечения двух окружностей, имеющих в данной системе координат XOY центры O' (1, 2) и O'' (1, 4) и радиусы $R_1=5$, $R_2=\sqrt{13}$.

Составим уравнения этих окружностей:

$$\begin{cases} (x-1)^2 + (y-2)^2 = 25, \\ (x-1)^2 + (y-4)^2 = 13. \end{cases}$$

Решая систему из этих двух уравнений, получим два решения:

$$x_1=4, \quad y_1=6 \quad \text{и} \quad x_2=-2, \quad y_2=6.$$

Это значит, что данные окружности пересекаются в двух точках:

$$M_1(4, 6) \quad \text{и} \quad M_2(-2, 6).$$

3. Линия, определяемая уравнением. Пусть даны плоскость с выбранной на ней декартовой прямоугольной системой координат XOY и некоторое уравнение $F(x, y)=0$. Рассмотрим на плоскости все точки (x, y) , координаты которых удовлетворяют этому уравнению. Полученное геометрическое место точек можно использовать для геометрической иллюстрации соотношения $F(x, y)=0$.

Например, уравнению $y-x=0$ соответствует на плоскости XOY геометрическое место точек $M(x, y)$, у которых $x=y$. Этому условию удовлетворяют точки, лежащие на биссектрисе L первого и третьего координатного углов, и только такие точки. Следовательно, уравнение $y-x=0$ на плоскости XOY изображается прямой L .

Уравнению $y=x^2$, т. е. $y-x^2=0$, соответствует на плоскости геометрическое место точек $M(x, y)$, у которых $y=x^2$. Это геометрическое место точек представляет собой линию, называемую *параболой*. Ее свойства отмечались в элементарной математике и будут подробно рассматриваться ниже.

В рассмотренных двух примерах уравнениям вида $F(x, y)=0$ соответствовали на плоскости XOY определенные линии, причем данные уравнения $y-x=0$ и $y=x^2$ являлись уравнениями этих линий (прямой и параболы). Условимся и в общем случае называть линией, определяемой уравнением $F(x, y)=0$ (в выбранной системе координат), геометрическое место точек плоскости, координаты которых удовлетворяют этому уравнению. При этом само уравнение $F(x, y)=0$, в соответствии с ранее данным определением, называется уравнением такой линии.

Пример 1. Уравнению $y=0$ соответствует на плоскости XOY геометрическое место точек $M(x, y)$, у которых ординаты равны нулю. Такому условию удовлетворяют точки оси OX и только они. Следовательно, уравнение $y=0$ есть уравнение прямой — оси OX . Аналогично получаем, что $x=0$ — уравнение оси OY .

Пример 2. Уравнению $y+x=0$ соответствует на плоскости геометрическое место точек $M(x, y)$, у которых $y=-x$ (абсцисса и ордината равны по абсолютной величине, различаются по знаку). Такому условию удовлетворяют точки биссектрисы второго — четвертого координатного углов, и только они.

Таким образом, линия, определяемая уравнением $y+x=0$, есть прямая (рис. 28).

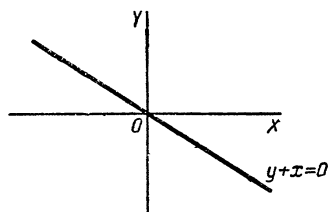


Рис. 28.

Пример 3. Уравнению $y^2-x^2=0$, т. е. $(y-x) \cdot (y+x)=0$, соответствуют на плоскости точки (x, y) , у которых $y-x=0$ или $y+x=0$. Учитывая рассмотренные выше примеры, заключаем, что линия, определяемая уравнением $y^2-x^2=0$, состоит из двух прямых — биссектрис первого — третьего и второго — четвертого координатных углов (рис. 29).

Пример 4. Уравнению $(x^2 - y^2) \cdot (x^2 + y^2 - 4) = 0$ соответствуют на плоскости точки (x, y) , у которых $x - y = 0$ или $x + y = 0$ или $x^2 + y^2 - 4 = 0$ ($x = y$, $x = -y$ или $x^2 + y^2 = 4$).

Таким образом, определяемая данным уравнением линия состоит из двух прямых и окружности (рис. 30)¹⁾.

Условимся в дальнейшем, как это принято в аналитической геометрии и других разделах высшей математики, вместо того чтобы говорить: «дана линия, определяемая уравнением $F(x, y) = 0$ », говорить: «дана линия (или кривая) $F(x, y) = 0$ ».

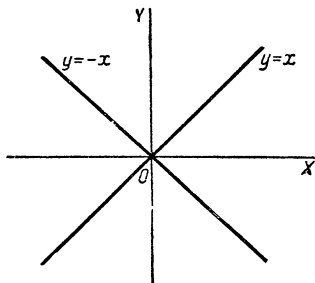


Рис. 29.

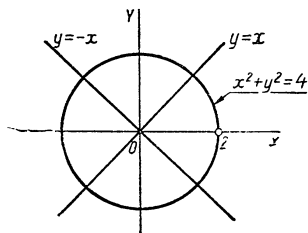


Рис. 30.

Примечания: 1. Следует заметить, что линия, определяемая уравнением $F(x, y) = 0$, может отличаться от того, что мы интуитивно привыкли считать линией²⁾. Так, например, уравнению $(x - 1)^2 + (y + 2)^2 = 0$ соответствует на плоскости одна точка (линия, определяемая этим уравнением, — точка), уравнению $(x^2 + y^2) \cdot (x^2 + y^2 - 1) = 0$ соответствует точка $(0, 0)$ и окружность с центром $(0, 0)$ радиуса 1 (линия — точка и окружность) и т. д.

2. Может оказаться, что не существует пар действительных чисел, удовлетворяющих уравнению $F(x, y) = 0$. Тогда на плоскости XOY ему не соответствует ни одной точки. Будем в этом случае говорить, что уравнение $F(x, y) = 0$ определяет *мнимую* линию.

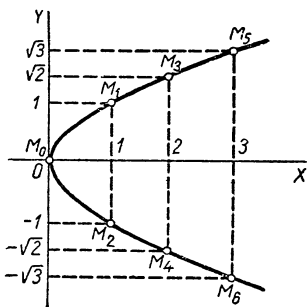


Рис. 31.

Например, уравнения $x^4 + y^4 + 4 = 0$, $\sin(x + y) - 3 = 0$ не удовлетворяются ни одной парой действительных чисел, т. е. определяют на плоскости XOY мнимые линии.

Приближенное построение линии, заданной уравнением $F(x, y) = 0$, можно производить «по точкам». Рассматриваем систему значений переменной x : $x_1, x_2, \dots, x_n$, отыскиваем для каждого такого значения x те значения переменной y , которые совместно с этим x удовлетворяют уравнению $F(x, y) = 0$, наносим на плоскость XOY точки с найденными координатами и соединяем их плавной линией. Чем чаще взяты значения x , тем, вообще говоря, точнее будет построена соответствующая линия.

Пример. Дано уравнение $y^2 - x = 0$, т. е. $y^2 = x$. Так как рассматриваются лишь пары действительных чисел, удовлетворяющих уравнению, то следует брать $x \geq 0$. Пусть $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, ...

Если $x = x_1 = 0$, то $y_1 = 0$.

Если $x = x_2 = 1$, то $y_2^2 = 1$, $y_2' = 1$, $y_2'' = -1$.

Если $x = x_3 = 2$, то $y_3^2 = 2$, т. е. $y_3' = \sqrt{2}$, $y_3'' = -\sqrt{2}$ и т. д.

Наносим на плоскость XOY точки $M_0(0, 0)$, $M_1(1, 1)$, $M_2(1, -1)$, $M_3(2, \sqrt{2})$, $M_4(2, -\sqrt{2})$ и т. д. и соединяем их плавной линией (рис. 31).

¹⁾ В этом случае говорят, что данная кривая распадается на три кривые.

²⁾ Строгое определение линии дается в специальной математической литературе.

Пусть теперь на плоскости введена полярная система координат.

Линией, определяемой уравнением $F(\rho, \varphi) = 0$, называется геометрическое место точек, полярные координаты которых (в выбранной системе координат) удовлетворяют уравнению $F(\rho, \varphi) = 0$.

Рассмотрим несколько примеров построения линий, заданных своими уравнениями в полярной системе координат, или, короче, своими полярными уравнениями.

Пример 1. Линия, определяемая уравнением $\rho = a\varphi$, где a — положительное число, называется *спиралью Архимеда*.

Для построения этой линии будем давать полярному углу φ различные неотрицательные значения и определять соответствующие значения полярного радиуса ρ . В таблице указаны значения угла φ , составляющие арифметическую прогрессию с разностью $\frac{\pi}{2}$, и соответствующие им значения ρ (они также составляют арифметическую прогрессию с разностью, равной $\frac{a\pi}{2}$).

φ	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π	$\frac{5\pi}{2}$	3π	...
ρ	0	$\frac{a\pi}{2}$	$a\pi$	$\frac{3a\pi}{2}$	$2a\pi$	$\frac{5a\pi}{2}$	$3a\pi$	...

Нанося на плоскость точки с найденными полярными координатами (ρ, φ) и соединяя их плавной линией, получим приближенно спираль Архимеда (рис. 32).

Исследуя уравнение $\rho = a\varphi$, $a > 0$, можно получить некоторые общие соображения, помогающие точнее построить спираль Архимеда. Так, например, поскольку уравнение $\rho = a\varphi$ удовлетворяется парой чисел $(0, 0)$, заключаем, что спираль Архимеда проходит через полюс O . Далее, если угол φ возрастает, то пропорционально ему возрастает ρ .

Т. е., проводя плавную линию через намеченные выше точки плоскости, мы

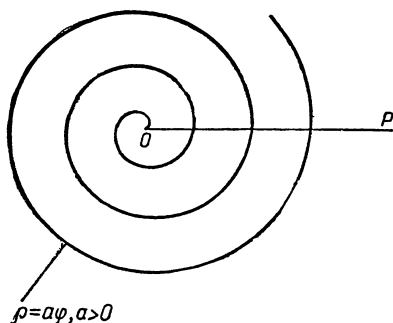


Рис. 32.

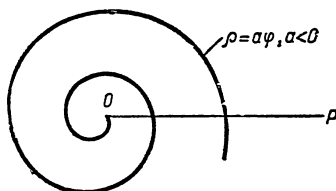


Рис. 33.

должны вести ее так, чтобы большим значениям φ соответствовали большие значения ρ . Далее, из уравнения $\rho = a\varphi$ следует, что для значений φ :

$$\varphi = \varphi_0, \quad \varphi = \varphi_0 + 2\pi, \quad \varphi = \varphi_0 + 4\pi, \quad \varphi = \varphi_0 + 6\pi, \quad \text{и т. д.}$$

соответствующие значения полярного радиуса составляют арифметическую прогрессию:

$$\rho = a\varphi_0, \quad \rho = a\varphi_0 + 2\pi a, \quad \rho = (a\varphi_0 + 2\pi a) + 2\pi a, \quad \rho = (a\varphi_0 + 4\pi a) + 2\pi a, \quad \dots$$

Это значит, что каждый луч, выходящий из полюса, спираль Архимеда пересекает на равные отрезки длины $2\pi a$.

Пример 2. Линия $\rho = a\varphi$, где $a < 0$, также называется *спиралью Архимеда*. Построение производится аналогично построению в предыдущем примере, но теперь для получения неотрицательных значений ρ приходится рассматривать $\varphi \leq 0$ (рис. 33).

Пример 3. Линия $\rho = \frac{a}{\varphi}$, где $a > 0$, называется *гиперболической спиралью*. Для ее построения составим таблицу значений ρ и φ , $\varphi > 0$,

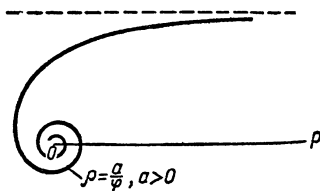


Рис. 34.

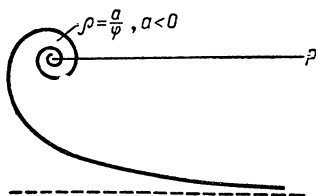


Рис. 35.

удовлетворяющих уравнению $\rho = \frac{a}{\varphi}$, и тем самым определим ряд точек этой линии.

φ	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{3\pi}{2}$	2π	$\frac{5\pi}{2}$	...
ρ	$\frac{6a}{\pi}$	$\frac{4a}{\pi}$	$\frac{3a}{\pi}$	$\frac{2a}{\pi}$	$\frac{2a}{3\pi}$	$\frac{a}{2\pi}$	$\frac{2a}{5\pi}$	...

Учитывая, что при неограниченном возрастании φ полярный радиус $\rho = \frac{a}{\varphi}$ неограниченно убывает, заключаем, что при этом соответствующая точка M кривой неограниченно приближается к полюсу O . При стремле-

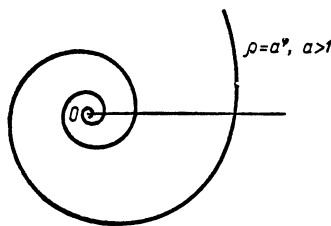


Рис. 36.

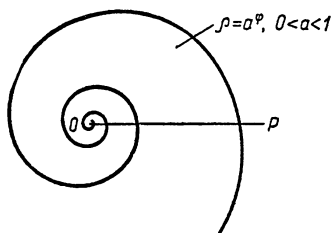


Рис. 37.

нии φ к нулю полярный радиус ρ неограниченно возрастает, т. е. соответствующая точка кривой неограниченно удаляется от полюса. Можно показать, что при этом точка M неограниченно приближается к прямой, параллельной полярной оси и отстоящей от нее на расстояние a .

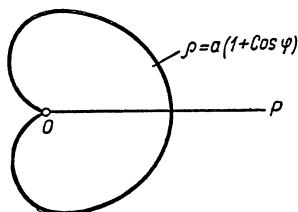


Рис. 38.

Нанося на плоскость найденные выше точки, получим приближенно гиперболическую спираль (рис. 34).

Аналогичным образом производится построение кривых $\rho = \frac{a}{\varphi}$, $a < 0$, $\varphi < 0$, (*гиперболическая спираль*), $\rho = a^\varphi$, $a > 1$ и $0 < a < 1$ (*логарифмические спирали*), $\rho = a(1 + \cos \varphi)$, $a > 0$ (*кардиоиды*) и др. (см. рис. 35, 36, 37, 38).

Алгебраические линии. Уравнение $F(x, y) = 0$ называется *алгебраическим*, если выражение $F(x, y)$ есть сумма конечного числа слагаемых вида Ax^ky^m , где k и m — целые неотрицательные числа, A — действительные числа. При этом наибольшая из сумм степеней $k + m$ называется *степенью уравнения*.

Так, например, уравнения

$$x + y - 5 = 0, \quad xy + 7 - x + 3y = 0, \quad x^2y^5 - x + y + 9 = 0$$

и т. д. будут алгебраическими. Степени их соответственно равны 1, 2, 8.

Уравнения $x + 2y - 1 = 0$, $\sin(x - y) + x - y = 1$ и т. д. алгебраическими не являются.

Общий вид алгебраического уравнения первой степени:

$$Ax + By + C = 0,$$

где A , B и C — некоторые числа, называемые *коэффициентами* уравнения.

Общий вид алгебраического уравнения второй степени:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0,$$

A , B , C , D , E , F — также числа, называемые *коэффициентами* уравнения, и т. д.

В аналитической геометрии обычно подробно изучаются *алгебраические линии первого и второго порядков*, т. е. линии, уравнения которых в некоторой декартовой прямоугольной системе координат являются алгебраическими уравнениями соответственно первой и второй степени.

Примечание. Уравнение алгебраической линии второго порядка в декартовых координатах удобно записывать в виде:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0.$$

Это упростит в дальнейшем запись ряда преобразований.

Раздел 3

ПРЯМАЯ ЛИНИЯ

§ 10. Угловой коэффициент

В настоящем разделе будут рассматриваться прямые линии и соответствующие им уравнения.

Эти прямые будем считать лежащими в плоскости, в которой выбрана определенная декартова прямоугольная система координат XOY .

Возьмем произвольную прямую. *Углом наклона* этой прямой к оси абсцисс называется угол, на который надо повернуть вокруг начала

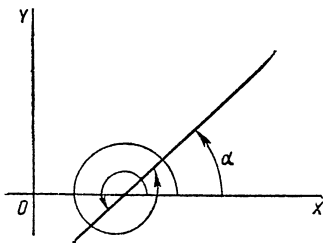


Рис. 39.

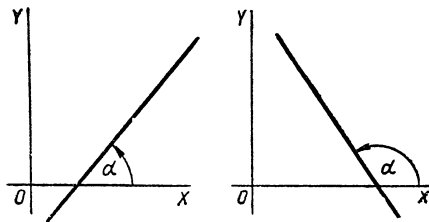


Рис. 40.

ось абсцисс, так чтобы она стала параллельна данной прямой. Этот угол считается положительным, если поворот оси абсцисс совершается против часовой стрелки, и отрицательным — в противном случае. Очевидно, что угол наклона прямой к оси абсцисс определяется неоднозначно (если α — угол наклона прямой к оси абсцисс, то $\alpha + n\pi$, где n — любое целое число, также будет углом наклона (рис. 39). Тангенсы всех этих углов наклона равны между собой.

Тангенс угла наклона α прямой к оси абсцисс называется *угловым коэффициентом* этой прямой линии и обозначается буквой k : $\operatorname{tg} \alpha = k$.

Если угол наклона прямой к оси абсцисс острый, то $k > 0$, если тупой, то $k < 0$ (рис. 40). Если прямая параллельна оси абсцисс, то $k = 0$. Если прямая перпендикулярна к оси абсцисс, то об угловом коэффициенте прямой говорить не имеет смысла. Так как при α , неограниченно приближающемся к $\frac{\pi}{2}$, тангенс угла наклона неограниченно растет, то иногда говорят, что у прямой, параллельной оси OY , угловой коэффициент равен бесконечности.

§ 11. Уравнение прямой линии с заданным угловым коэффициентом

Требуется составить уравнение прямой, не параллельной оси ординат, зная величину b отрезка $\overline{OB}$, отсекаемого этой прямой на оси ординат, и угловой коэффициент k .

Пусть l — данная прямая (рис. 41). Установив на ней какое-нибудь направление, обозначим через α угол между осью OX и осью l . Для любой точки $M(x, y)$ на рассматриваемой прямой имеем:

$$\operatorname{pr}_{OX} \overline{BM} = \overline{ON} = x,$$

$$\operatorname{pr}_{OY} \overline{BM} = \overline{BL} = \overline{BO} + \overline{OL} = -b + y = y - b,$$

а с другой стороны (см. сноску на стр. 21),

$$\operatorname{pr}_{OX} \overline{BM} = \overline{BM} \cos \alpha,$$

$$\operatorname{pr}_{OY} \overline{BM} = \overline{BM} \cos \left(\alpha - \frac{\pi}{2} \right) = \overline{BM} \sin \alpha.$$

Следовательно,

$$y - b = \overline{BM} \sin \alpha = \underbrace{\overline{BM} \cos \alpha}_{x} \operatorname{tg} \alpha = x \operatorname{tg} \alpha$$

или

$$y = kx + b. \quad (1)$$

Уравнению (1) удовлетворяют координаты любой точки, лежащей на прямой, и не удовлетворяют координаты точек, не лежащих на данной прямой. Уравнение (1) называется *уравнением прямой с угловым коэффициентом*. В этом уравнении x и y являются координатами произвольной точки прямой («текущими координатами»), а постоянные для данного уравнения величины b и k называются параметрами уравнения: k — *угловой коэффициент*, а b — *начальная ордината*.

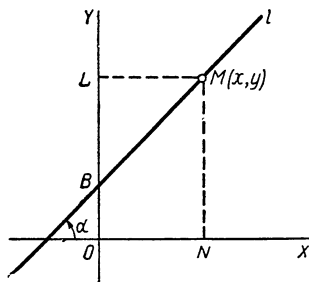


Рис. 41.

Геометрический смысл этих параметров заключается в следующем: угловой коэффициент k равен тангенсу угла наклона прямой к оси абсцисс; начальная ордината b есть величина отрезка, отсекаемого прямой на оси ординат.

Рассмотрим частные случаи уравнения прямой с угловым коэффициентом $y = kx + b$.

1) $b = 0$, уравнение имеет вид:

$$y = kx. \quad (2)$$

Соответствующая прямая проходит через начало координат.

2) $k = 0$, уравнение прямой имеет вид:

$$y = b. \quad (3)$$

Прямая (3) параллельна оси абсцисс и отсекает на оси OY отрезок, величина которого равна b .

Пример. У прямой, заданной уравнением $y = -x + 3$, начальная ордината $b = 3$, а угловой коэффициент $k = \operatorname{tg} \alpha = -1$, откуда (беря наименьшее положительное значение угла) получаем: $\alpha = \frac{3\pi}{4}$. Следовательно, прямая наклонена к оси

абсцисс под углом $\alpha = \frac{3\pi}{4}$ и отсекает на оси ординат отрезок, величина которого равна 3.

Если прямая параллельна оси OY и отсекает на оси OX отрезок, величина которого равна a , то координаты текущей точки $M(x, y)$ прямой удовлетворяет уравнению

$$x = a. \quad (4)$$

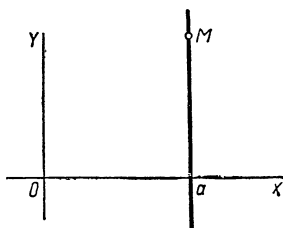


Рис. 42.

Это уравнение и будет уравнением данной прямой (рис. 42).

Рассматривая полученные нами уравнения (1), (4), убеждаемся, что во всех случаях прямая может быть задана уравнением первой степени относительно переменных x и y .

§ 12. Общее уравнение прямой и его исследование

Докажем теперь обратное утверждение, т. е. докажем, что всякое уравнение первой степени относительно x и y :

$$Ax + By + C = 0^1 \quad (1)$$

определяет в прямоугольной системе координат XOY некоторую прямую.

Рассмотрим возможные случаи.

1) $B \neq 0$; $A \neq 0$; $C \neq 0$. Разделив все члены уравнения на B и определяя на него y , запишем уравнение (1) в виде: $y = -\frac{A}{B}x - \frac{C}{B}$.

Обозначая

$$-\frac{A}{B} = k, \quad (2)$$

$$-\frac{C}{B} = b, \quad (3)$$

получим уравнение $y = kx + b$. Это уравнение прямой, угловой коэффициент и начальная ордината которой определяются равенствами (2) и (3).

2) $B \neq 0$; $C \neq 0$; $A = 0$. В этом случае уравнение принимает вид $By + C = 0$; $y = -\frac{C}{B}$. Это уравнение прямой, параллельной оси абсцисс и отсекающей от оси ординат отрезок, величина которого равна b , где $b = -\frac{C}{B}$.

3) $A \neq 0$; $C \neq 0$; $B = 0$. Уравнение (1) принимает вид: $Ax + C = 0$; $x = -\frac{C}{A}$. Полагая $a = -\frac{C}{A}$, полученное уравнение запишем в виде $x = a$. Это уравнение прямой, параллельной оси ординат и отсекающей от оси абсцисс отрезок, величина которого a .

4) $A \neq 0$; $B \neq 0$; $C = 0$. Уравнение (1) имеет вид: $Ax + By = 0$; $y = -\frac{A}{B}x$; или $y = kx$, где $k = -\frac{A}{B}$. Это уравнение прямой, проходящей через начало координат.

¹⁾ Предполагается, что коэффициенты A и B одновременно не равны нулю.

5) $B \neq 0$; $C = 0$; $A = 0$. Уравнение (1) имеет вид:

$$y = 0. \quad (4)$$

Это уравнение является уравнением оси абсцисс.

6) $A \neq 0$; $B = 0$; $C = 0$. Уравнение имеет вид:

$$x = 0. \quad (5)$$

Это — уравнение оси ординат.

Таким образом, во всех случаях уравнение $Ax + By + C = 0$ является уравнением прямой линии. Уравнение $Ax + By + C = 0$, где A и B одновременно не равны нулю, называется *общим уравнением прямой*. Итак, в прямоугольной декартовой системе координат всякая прямая может быть представлена уравнением первой степени, и, наоборот, любое уравнение первой степени относительно x и y определяет прямую линию.

§ 13. Уравнение прямой проходящей через данную точку в данном направлении; уравнение прямой, проходящей через две данные точки; уравнение прямой в отрезках на осях

Решим несколько задач на составление уравнения прямой линии по различным данным.

Задача 1. Составить уравнение прямой, проходящей через точку $M_0(x_0, y_0)$ и имеющей угловой коэффициент k .

Решение. Уравнение этой прямой можно записать в виде:

$$y = kx + b, \quad (1)$$

где b — пока неизвестная начальная ордината.

Так как искомая прямая проходит через точку $M_0(x_0, y_0)$, то координаты этой точки должны удовлетворять уравнению (1):

$$y_0 = kx_0 + b. \quad (2)$$

Определяя b из равенства (2) и подставляя его в уравнение (1), получим искомое уравнение прямой:

$$y - y_0 = k(x - x_0). \quad (3)$$

Уравнение (3) называется *уравнением прямой, проходящей через данную точку в данном направлении*.

Задача 2. Составить уравнение не вертикальной прямой, проходящей через точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$.

Решение. Уравнение этой прямой можно записать в виде:

$$y - y_1 = k(x - x_1), \quad (4)$$

где k — пока неизвестный угловой коэффициент.

Так как искомая прямая проходит через точку $M_2(x_2, y_2)$, то координаты этой точки должны удовлетворять уравнению (4):

$$y_2 - y_1 = k(x_2 - x_1). \quad (5)$$

Отсюда находим

$$k = \frac{y_2 - y_1}{x_2 - x_1}. \quad (6)$$

Подставляя найденное значение k в уравнение (4), после преобразования получим искомое уравнение прямой, проходящей через точки M_1 и M_2 . Если $y_1 \neq y_2$, то его можно записать в виде:

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \quad (7)$$

Если $y_1 = y_2$, то уравнение искомой прямой $y = y_1$.

Уравнение (7) называется *уравнением прямой, проходящей через две точки*. Заметим, что равенство (6) позволяет определить угловой коэффициент неперпендикулярной прямой по двум ее точкам.

Если $x_1 = x_2$, то прямая, проходящая через точки M_1 и M_2 , параллельна оси OY , ее уравнение имеет вид: $x = x_1$.

Задача 3. Составить уравнение прямой, отсекающей на оси абсцисс отрезок величины a ($a \neq 0$), а на оси ординат — отрезок величины b , $b \neq 0$.

Решение. Пусть данная прямая отсекает на оси абсцисс отрезок $\overline{OA}$, а на оси ординат отрезок $\overline{OB}$. Тогда координаты точки A будут $(a, 0)$, а координаты точки B будут $(0, b)$ (рис. 43).

Составим уравнение данной прямой, как прямой, проходящей через две точки $A(a, 0)$ и $B(0, b)$.

Подставляя координаты точек в (7), получим уравнение

$$\frac{y}{b} = \frac{x-a}{-a}$$

или

$$\frac{x}{a} + \frac{y}{b} = 1. \quad (8)$$

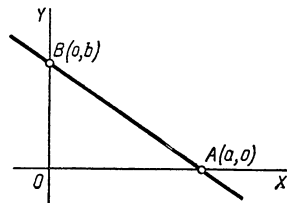


Рис. 43.

Уравнение (8) называется *уравнением прямой в отрезках*. В этом уравнении x и y — текущие координаты, a и b — параметры. Заметим, что это уравнение удобно использовать для геометрического построения прямой. Откладывая на оси абсцисс отрезок, равный по величине a , на оси ординат отрезок, равный по величине b , и соединяя концы этих отрезков, получим искомую прямую.

§ 14. Угол между двумя прямыми; условия параллельности и перпендикулярности двух прямых

Возьмем две прямые, не параллельные оси ординат: уравнение прямой L_1 пусть будет $y = k_1x + b_1$, где $k_1 = \operatorname{tg} \alpha_1$, а уравнение прямой L_2 $y = k_2x + b_2$, где $k_2 = \operatorname{tg} \alpha_2$, (рис. 44).

Определим угол наклона φ прямой L_2 к прямой L_1 , т. е. определим угол, на который надо повернуть прямую L_1 , чтобы она стала параллельной прямой L_2 . Как и при определении угла между осями, угол φ считаем положительным в случае вращения прямой L_1 против часовой стрелки и отрицательным — в противном случае.

На основании геометрических соображений устанавливаем зависимость между углами α_1 , α_2 , φ :

$$\alpha_2 = \varphi + \alpha_1, \quad \varphi = \alpha_2 - \alpha_1.$$

Если $\varphi \neq \frac{\pi}{2}$, то

$$\begin{aligned} \operatorname{tg} \varphi &= \operatorname{tg} (\alpha_2 - \alpha_1), \\ \operatorname{tg} \varphi &= \frac{\operatorname{tg} \alpha_2 - \operatorname{tg} \alpha_1}{1 + \operatorname{tg} \alpha_2 \operatorname{tg} \alpha_1}, \end{aligned}$$

а значит,

$$\operatorname{tg} \varphi = \frac{k_2 - k_1}{1 + k_1 k_2}. \quad (1)$$

Формула (1) служит для определения *угла между двумя неперпендикулярными прямыми*.

Если прямые L_1 и L_2 параллельны (и, в частности, совпадают), то $\varphi = 0$ и $\operatorname{tg} \varphi = 0$. В этом случае в формуле (1) числитель в правой части должен быть равен нулю:

$$k_2 - k_1 = 0$$

т. е.

$$k_2 = k_1 \quad (2)$$

и обратно, если прямые L_1 и L_2 таковы, что $k_1 = k_2$, то $\operatorname{tg} \varphi = 0$, т. е. эти прямые параллельны.

Таким образом, равенство (2) является *условием параллельности* двух (невертикальных) *прямых*. Это условие параллельности прямых может быть сформулировано так: если прямые параллельны, то их угловые коэффициенты равны, и наоборот.

Если прямые L_1 и L_2 перпендикулярны,

т. е. $\varphi = \frac{\pi}{2}$, то

$$\alpha_2 = \frac{\pi}{2} + \alpha_1, \quad \operatorname{tg} \alpha_2 = -\operatorname{ctg} \alpha_1,$$

т. е.

$$k_2 = -\frac{1}{k_1}. \quad (3)$$

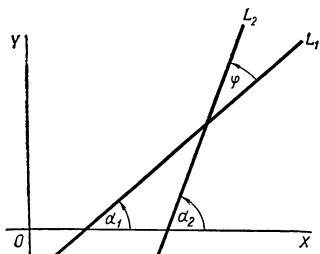


Рис. 44.

Наоборот, если для прямых L_1 и L_2 имеем $k_2 = -\frac{1}{k_1}$, то $\operatorname{tg} \alpha_2 = \operatorname{tg} \left(\frac{\pi}{2} + \alpha_1 \right)$, т. е. $\varphi = \frac{\pi}{2}$, прямые L_1 и L_2 перпендикулярны.

Равенство (3) является *условием перпендикулярности* двух (невертикальных) *прямых*. Таким образом, условие перпендикулярности может быть сформулировано так: если прямые перпендикулярны, то их угловые коэффициенты обратны по величине и противоположны по знаку, и наоборот.

Если уравнения прямых L_1 и L_2 даются в общем виде:

$$A_1x + B_1y + C_1 = 0 \quad \text{и} \quad A_2x + B_2y + C_2 = 0,$$

то на основании равенства (2) (§ 12) имеем:

$$k_1 = -\frac{A_1}{B_1}; \quad k_2 = -\frac{A_2}{B_2}. \quad ^1)$$

Подставляя значения k_1 и k_2 в равенство (1), получим:

$$\operatorname{tg} \varphi = \frac{A_1B_2 - A_2B_1}{A_1A_2 + B_1B_2}. \quad (4)$$

В этом случае условие параллельности прямых записывается в виде:

$$A_1B_2 - A_2B_1 = 0, \quad (5)$$

а условие перпендикулярности прямых — в виде:

$$A_1A_2 + B_1B_2 = 0. \quad (6)$$

Примечание. Как легко видеть, условия (5) и (6) остаются в силе и для прямых, параллельных оси ординат.

¹⁾ Так как по условию прямые не параллельны оси ординат, то $B_1 \neq 0$; $B_2 \neq 0$.

§ 15. Пересечение двух прямых

Прямые L_1 и L_2 заданы уравнениями:

$$\begin{aligned} A_1x + B_1y + C_1 &= 0, \\ A_2x + B_2y + C_2 &= 0. \end{aligned} \quad (1)$$

Требуется определить точку пересечения этих прямых.

Пусть $A_1B_2 - A_2B_1 \neq 0$. Тогда система (1) имеет единственное решение¹⁾:

$$x = \frac{B_1C_2 - B_2C_1}{A_1B_2 - A_2B_1}, \quad y = \frac{A_2C_1 - A_1C_2}{A_1B_2 - A_2B_1}. \quad (2)$$

Это значит, что прямые L_1 и L_2 не параллельны, пересекаются в одной точке. Ее координаты даются формулами (2).

Пусть теперь $A_1B_2 - A_2B_1 = 0$. При этом возможны два случая. Либо, кроме условия $A_1B_2 - A_2B_1 = 0$, выполнено также условие $A_1C_2 - A_2C_1 = 0$ (а тогда и $B_1C_2 - B_2C_1 = 0$), либо $A_1B_2 - A_2B_1 = 0$, но $A_1C_2 - A_2C_1 \neq 0$ (или $B_1C_2 - B_2C_1 \neq 0$).

В первом случае имеем $A_2 = \mu A_1$, $B_2 = \mu B_1$, $C_2 = \mu C_1$, откуда следует, что второе уравнение системы получается из первого умножением на число μ . При этом всякая пара чисел x, y , удовлетворяющая первому уравнению, удовлетворяет и второму, т. е. всякая точка прямой L_1 принадлежит и прямой L_2 . Это значит, что прямые L_1 и L_2 совпадают.

Равенства

$$A_2 = \mu A_1, \quad B_2 = \mu B_1, \quad C_2 = \mu C_1, \quad (3)$$

где $\mu \neq 0$ — некоторое число, обычно записывают в виде

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} \quad (4)$$

и называют *условиями совпадения прямых*.

Эти равенства (4) имеют смысл, если знаменатели отношений не равны нулю. В случае, если в каком-либо отношении знаменатель окажется равным нулю, то [см. (3)] равен нулю и соответствующий числитель. Можно договориться условия совпадения снова писать в виде (4), вычеркивая из этих отношений то, у которого знаменатель равен нулю, и приравнивая нулю его числитель.

Если же $A_1B_2 - A_2B_1 = 0$, но, например, $A_1C_2 - A_2C_1 \neq 0$, то, допустив, что система (1) имеет решение (x_0, y_0) , мы получаем противоречие.

Действительно, подставив в уравнения вместо x и y значения x_0 и y_0 , умножив первое равенство на A_2 , второе на A_1 и вычитая из первого второе, получим $C_1A_2 - C_2A_1 = 0$, что противоречит предположению.

Таким образом, в рассматриваемом случае система решения не имеет. Это значит, что прямые L_1 и L_2 не имеют ни одной точки пересечения, т. е. они параллельны.

Таким образом, условия параллельности прямых:

$$\begin{aligned} A_1B_2 - A_2B_1 &= 0, & A_1C_2 - A_2C_1 &\neq 0 \\ \text{или } B_1C_2 - B_2C_1 &\neq 0. \end{aligned}$$

Условие параллельности прямых обычно записывают в виде:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2}.$$

(При этом, если у отношения $\frac{A_1}{A_2}$ или $\frac{B_1}{B_2}$ знаменатель равен нулю, то равен нулю и числитель, и соответствующее отношение надо отбросить.)

¹⁾ Решение такой системы подробно рассматривается в § 34 разд. 6.

Резюмируя, можно сказать, что прямые на плоскости либо пересекаются в одной точке, либо параллельны, либо совпадают.

Примечание. Если совпадающие прямые рассматривать как частный случай параллельных прямых, то условие параллельности прямых L_1 и L_2 будет иметь вид:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2}.$$

§ 16. Нормальное уравнение прямой

Выведем сначала уравнение произвольной прямой L в полярных координатах.

Пусть l — ось, проходящая через начало O и перпендикулярная данной прямой L (направление l берется от O в сторону данной прямой¹⁾), p — расстояние от O до данной прямой, α — угол между полярной осью OP и осью l (рис. 45).

Для любой точки $M(p, \varphi)$ на данной прямой L имеем:

$$\text{пр}_l \overrightarrow{OM} = \overrightarrow{ON} = p,$$

но, с другой стороны,

$$\text{пр}_l \overrightarrow{OM} = OM \cos(\varphi - \alpha) = p \cos(\varphi - \alpha),$$

следовательно,

$$p \cos(\varphi - \alpha) = p. \quad (1)$$

Так как точки, не лежащие на данной прямой L , этому соотношению не удовлетворяют, то (1) есть уравнение прямой L в полярных координатах.

Рассмотрим прямоугольную систему координат XOY и прямую L . Введя полярные координаты (взяв O за полюс и OX за полярную ось), получим уравнение (1) прямой L в полярных координатах. Это уравнение можно переписать в виде:

$$p \cos \varphi \cos \alpha + p \sin \varphi \sin \alpha - p = 0.$$

В силу формул, связывающих прямоугольную систему координат с полярной, имеем $p \cos \varphi = x$ и $p \sin \varphi = y$. Следовательно уравнение (1) прямой L в прямоугольной системе координат примет вид:

$$x \cos \alpha + y \sin \alpha - p = 0. \quad (2)$$

Уравнение (2) называется *нормальным уравнением прямой* в прямоугольной системе координат.

В полученном нормальном уравнении прямой L числа p и α , где p — длина перпендикуляра, опущенного из начала координат на прямую, а α — угол наклона нормали l к оси абсцисс, являются параметрами уравнения.

Покажем, как привести уравнение прямой в общем виде к уравнению той же прямой в нормальном виде.

Возьмем уравнение прямой L

$$Ax + By + C = 0 \quad (3)$$

и уравнение (2) той же прямой в нормальном виде

$$x \cos \alpha + y \sin \alpha - p = 0.$$

¹⁾ Если данная прямая проходит через точку O , то выбор направления оси l произволен.

После умножения уравнения (3) на множитель λ , $\lambda \neq 0$, получим:

$$\lambda Ax + \lambda By + \lambda C = 0. \quad (4)$$

При надлежащем подборе λ уравнение (3) должно обратиться в уравнение (2). Следовательно, должны выполняться равенства: $\lambda A = \cos \alpha$; $\lambda B = \sin \alpha$; $\lambda C = -p$. Из первых двух равенств определим λ , для чего каждое из них возведем в квадрат, а затем сложим:

$$\lambda^2 A^2 + \lambda^2 B^2 = \cos^2 \alpha + \sin^2 \alpha = 1.$$

Получим:

$$\lambda = \frac{1}{\pm \sqrt{A^2 + B^2}}. \quad (5)$$

Для выбора в этой формуле знака перед квадратным корнем используем равенство $\lambda C = -p$. Если прямая L не проходит через начало координат, то p , как длина перпендикуляра, есть число положительное, следовательно, произведение $\lambda C < 0$, а для этого необходимо, чтобы λ и C имели противоположные знаки. Следовательно, в формуле (5) знак перед корнем берется противоположный знаку C . Если же прямая L проходит через начало координат, то $C = 0$, выбор знака λ произволен.

Множитель λ называется *нормирующим множителем*. Итак, для приведения общего уравнения прямой к нормальному виду надо найти значение λ , а затем все члены уравнения умножить на λ .

§ 17. Расстояние от точки до прямой

С помощью нормального уравнения прямой можно определить *расстояние от данной точки до данной прямой*.

Пусть дается некоторая прямая L своим уравнением в нормальном виде $x \cos \alpha + y \sin \alpha - p = 0$ и точка $M_0(x_0, y_0)$, лежащая вне прямой. Требуется определить расстояние d от точки M_0 до этой прямой.

Через точку M_0 проведем прямую L_0 параллельно прямой L и рассмотрим ее нормальное уравнение

$$x \cos \alpha_0 + y \sin \alpha_0 - p_0 = 0.$$

Так как точка $M_0(x_0, y_0)$ лежит на прямой L_0 , то

$$x_0 \cos \alpha_0 + y_0 \sin \alpha_0 - p_0 = 0. \quad (1)$$

В зависимости от положения прямой L и точки M_0 будем иметь:

$p_0 = p + d$; $\alpha_0 = \alpha$, если точки M_0 и O лежат по разные стороны прямой L (рис. 46),

$p_0 = p - d$; $\alpha_0 = \alpha$, если точки M_0 и O лежат по одну сторону прямой L и $p \geq d$ (рис. 47),

$p_0 = d - p$; $\alpha_0 = \alpha + \pi$, если точки M_0 и O лежат по одну сторону прямой L и $p \leq d$ (рис. 48).

Подставляя эти выражения для p_0 и α_0 в (1), получим:

$$\pm d = x_0 \cos \alpha + y_0 \sin \alpha - p, \quad (2)$$

где в левой части берется знак плюс, если точки M_0 и O лежат по разные стороны прямой L , и знак минус, если M_0 и O лежат по одну сторону прямой L .

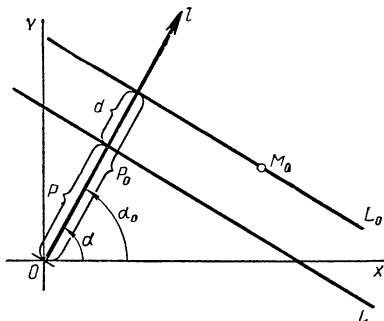


Рис. 46.

Отсюда следует, что для любого взаимного расположения точки M_0 и прямой L

$$d = |x_0 \cos \alpha + y_0 \sin \alpha - p|. \quad (3)$$

Примечание. Назовем *отклонением* δ точки M_0 от прямой L число $+d$, если точка M_0 и начало O лежат по разные стороны от прямой L ($\delta = d$), и число $-d$, если они лежат по одну сторону от L ($\delta = -d$). Тогда равенство (2) можно переписать так:

$$\delta = x_0 \cos \alpha + y_0 \sin \alpha - p.$$

Знак отклонения характеризует взаимное расположение точки M_0 , прямой L и начала O .

Итак, чтобы найти расстояние от данной точки до данной прямой, надо уравнение прямой привести к нормальному виду, вместо текущих координат подставить в левую часть уравнения координаты данной точки и взять абсолютную величину полученного результата.

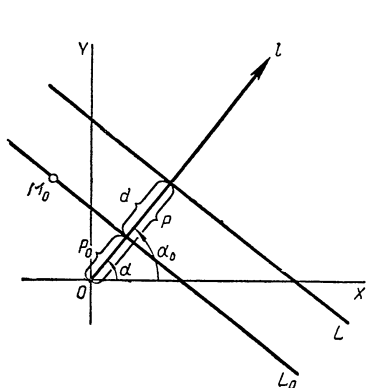


Рис. 47.

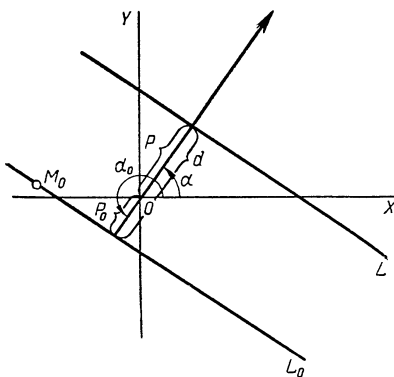


Рис. 48.

Если уравнение прямой дается в общем виде, то для решения задачи о нахождении расстояния от точки $M_0(x_0, y_0)$ до прямой $Ax + By + C = 0$ можно пользоваться формулой

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}. \quad (4)$$

Рекомендуем читателям, используя формулу (4), доказать, что биссектрисы углов, образованных прямыми

$$A_1x + B_1y + C_1 = 0, \quad A_2x + B_2y + C_2 = 0,$$

имеют следующие уравнения:

$$\frac{A_1x + B_1y + C_1}{\sqrt{A_1^2 + B_1^2}} = \frac{A_2x + B_2y + C_2}{\sqrt{A_2^2 + B_2^2}}, \quad (5)$$

$$\frac{A_1x + B_1y + C_1}{\sqrt{A_1^2 + B_1^2}} = -\frac{A_2x + B_2y + C_2}{\sqrt{A_2^2 + B_2^2}}. \quad (6)$$

§ 18. Уравнение пучка прямых

Возьмем на плоскости точку $M_0(x_0, y_0)$. Через эту точку можно провести бесчисленное множество прямых. Совокупность всех прямых плоскости, проходящих через данную точку, называется *пучком* прямых (рис. 49). Точка M_0 называется *центром пучка*. Любые две фиксированные прямые, принадлежащие пучку, называются *образующими* пучка прямых.

Поставим задачу: имея уравнения двух образующих пучка L_1 и L_2 :

$$A_1x + B_1y + C_1 = 0, \quad A_2x + B_2y + C_2 = 0,$$

найти уравнения других прямых пучка.

Составим уравнение

$$A_1x + B_1y + C_1 + \lambda(A_2x + B_2y + C_2) = 0, \quad (1)$$

где λ — параметр (произвольное число).

Уравнение (1), как уравнение первой степени, есть уравнение прямой при любом значении λ .

Так как прямые $A_1x + B_1y + C_1 = 0$ и $A_2x + B_2y + C_2 = 0$ проходят через точку $M_0(x_0, y_0)$, то $A_1x_0 + B_1y_0 + C_1 = 0$ и $A_2x_0 + B_2y_0 + C_2 = 0$.

Следовательно, координаты центра пучка $M_0(x_0, y_0)$ удовлетворяют уравнению (1) (при произвольном значении λ). Это значит, что при каждом значении λ уравнение (1) есть уравнение прямой данного пучка.

Чтобы выяснить, определяет ли уравнение (1) любую прямую, принадлежащую пучку, зададим произвольную точку $N(m, n)$ и среди прямых, имеющих уравнение (1), будем искать ту, которая проходит через эту точку.

Для этого в уравнение (1) вместо текущих координат подставим координаты точки $N(m, n)$ и найдем соответствующее значение λ :

$$\lambda = \frac{A_1m + B_1n + C_1}{A_2m + B_2n + C_2}. \quad (2)$$

Очевидно, что это значение λ можно найти всегда, кроме случая, когда $A_2m + B_2n + C_2 = 0$, т. е. когда заданная точка N лежит на образующей пучка — прямой

$$A_2x + B_2y + C_2 = 0.$$

Следовательно, уравнение (1) определяет все прямые пучка, кроме второй образующей.

Уравнение (1) называют *уравнением пучка прямых, заданного образующими*.

Уравнение всякой неперпендикулярной прямой, проходящей через точку $M_0(x_0, y_0)$, можно записать в виде:

$$y - y_0 = k(x - x_0), \quad (3)$$

где k — угловой коэффициент этой прямой.

Каждая (неперпендикулярная) прямая пучка с центром в точке M_0 имеет уравнение вида (3) (при соответствующем этой прямой значении параметра k). Наоборот, всякая прямая, имеющая уравнение (3), принадлежит рассматриваемому пучку. В связи с этим уравнение (3), где k — параметр, называют уравнением пучка прямых с центром $M_0(x_0, y_0)$.

Рассмотрим решения некоторых задач на прямую линию.

Задача 1. Зная уравнение прямой $3x - 4y - 60 = 0$, определить: а) отрезки a и b , которые прямая отсекает на координатных осях; б) угловой коэффициент к прямой; в) расстояние данной прямой от начала координат.

Решение. Для ответа на первый вопрос приведем уравнение прямой

$$3x - 4y - 60 = 0$$

к виду: $\frac{x}{20} - \frac{y}{15} = 1$. Отсюда следует, что $a = 20$, $b = -15$.

Приводя данное уравнение к виду $y = kx + b$, получим $y = \frac{3}{4}x - 15$, откуда находим $k = \frac{3}{4}$.

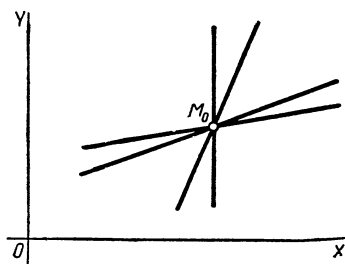


Рис. 49.

И, наконец, для ответа на последний вопрос приведем уравнение прямой к нормальному виду, для чего все члены уравнения $3x - 4y - 60 = 0$ умножим на нормирующий множитель λ , определяемый формулой

$$\lambda = \frac{1}{\pm \sqrt{9 + 16}}.$$

Так как свободный член отрицателен, то λ берется положительным, следовательно, $\lambda = \frac{1}{5}$.

Нормальное уравнение данной прямой:

$$\frac{3}{5}x - \frac{4}{5}y - 12 = 0.$$

Таким образом, $p = 12$, т. е. расстояние прямой от начала координат равно 12. **Задача 2.** В треугольнике ABC , заданном координатами своих вершин:

$$A(6, 4); \quad B(-9, 4); \quad C(-9, -16),$$

найти: а) длину медианы AE , проведенной из вершины A ; б) длину высоты BD , опущенной из вершины B ; в) точку пересечения медианы AE и высоты BD .

Сделаем схематический чертеж треугольника (рис. 50). Сначала определим координаты точки E , как точки, делящей отрезок $\overline{BC}$ пополам:

$$x_E = \frac{-9 - 9}{2} = -9, \quad y_E = \frac{4 - 16}{2} = -6, \\ \text{т. е. } E(-9, -6).$$

Длину медианы определим по формуле расстояния между двумя точками A и E :

$$AE = \sqrt{(6 + 9)^2 + (4 + 6)^2} = \sqrt{325}.$$

Далее составим уравнение стороны AC , используя уравнение прямой, проходящей через две точки A и C , а длину BD определим по формуле расстояния от точки B до прямой AC :

$$\frac{y + 16}{4 + 16} = \frac{x + 9}{6 + 9}, \quad 4x - 3y - 12 = 0, \quad (AC) \\ BD = \frac{|4 \cdot (-9) - 3 \cdot 4 - 12|}{\sqrt{16 + 9}} = 12.$$

Затем составим уравнение медианы AE , используя уравнение прямой, проходящей через две точки A и E :

$$\frac{y + 6}{4 + 6} = \frac{x + 9}{6 + 9}, \quad 2x - 3y = 0. \quad (AE)$$

Уравнение высоты BD составим по точке и угловому коэффициенту AC . Угловой коэффициент прямой AC : $k_{AC} = \frac{4}{3}$, следовательно, угловой коэффициент прямой BD , перпендикулярной к AC , равен

$$k_{BD} = -\frac{3}{4}.$$

Уравнение прямой BD имеет вид:

$$y - 4 = -\frac{3}{4}(x + 9), \quad 3x + 4y + 11 = 0. \quad (BD)$$

Для определения искомой точки M пересечения медианы и высоты решаем совместно уравнения прямых AE и BD . При решении системы находим, что

$$x_M = -\frac{33}{17}, \quad y_M = -\frac{22}{17}.$$

Задача 3. Треугольник задан уравнениями своих сторон

$$(AB) \dots x - 2y + 3 = 0; \quad (BC) \dots 5x + y - 7 = 0; \quad (AC) \dots 2x + 3y - 21 = 0.$$

Требуется найти точку, симметричную вершине B относительно основания AC .

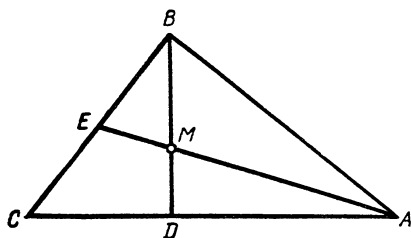


Рис. 50.

Решение. Сделаем схематический чертеж треугольника ABC (рис. 51). Из вершины B опустим перпендикуляр BD на сторону AC и продолжим эту прямую на расстояние $DM=BD$. Очевидно, точка M и будет искомой.

Для определения координат точки B решим совместно уравнения сторон AB и BC :

$$x - 2y + 3 = 0 \dots (AB), \quad 5x + y - 7 = 0 \dots (BC),$$

откуда $x_B = 1, y_B = 2$.

Составим уравнение прямой BD , как прямой, проходящей через точку B перпендикулярно стороне AC .

$$k_{AC} = -\frac{2}{3}, \text{ следовательно, } k_{BD} = \frac{3}{2},$$

$$y - 2 = \frac{3}{2}(x - 1) \dots (BD) \text{ или } 3x - 2y + 1 = 0.$$

Найдем координаты точки D , как точки пересечения прямых AC и BD :

$$4x + 6y - 42 = 0 \dots (AC), \quad 9x - 6y + 3 = 0 \dots (BD),$$

откуда $x_D = 3, y_D = 5$.

Для нахождения координат точки M используем формулы деления отрезка пополам, которые для данной задачи запишем так:

$$x_D = \frac{x_B + x_M}{2}, \quad y_D = \frac{y_B + y_M}{2}.$$

Подставляя в данные формулы известные координаты точек $B(1, 2)$ и $D(3, 5)$, найдем искомые координаты точки M :

$$x_M = 5; \quad y_M = 8.$$

Задача 4. Найти прямую, которая проходит через точку пересечения двух прямых $x + 2y - 19 = 0$; $2x - y - 3 = 0$ и находится от точки $M(2, 3)$ на расстоянии 5 единиц.

Решение. Используем так называемый *метод пучка прямых* (заданных своими образующими). За образующие можно взять прямые, данные в условии задачи. Уравнение пучка прямых будет иметь вид:

$$x + 2y - 19 + \lambda(2x - y - 3) = 0,$$

т. е.

$$(1 + 2\lambda)x + (2 - \lambda)y - 19 - 3\lambda = 0.$$

Чтобы определить значения параметра λ для искомой прямой, используем формулу расстояния от точки до прямой:

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}.$$

По условиям нашей задачи:

$$d = 5, A = 1 + 2\lambda, B = 2 - \lambda, C = -19 - 3\lambda, x_0 = 2, y_0 = 3.$$

Подставляя эти значения в приведенную формулу, получим следующее уравнение относительно параметра λ :

$$5 = \frac{|(1 + 2\lambda)2 + (2 - \lambda)3 - 19 - 3\lambda|}{\sqrt{(1 + 2\lambda)^2 + (2 - \lambda)^2}},$$

$$5\sqrt{(1 + 2\lambda)^2 + (2 - \lambda)^2} = |2\lambda + 11|.$$

Возводя в квадрат и приводя подобные члены, получим квадратное уравнение относительно λ : $121\lambda^2 - 44\lambda + 4 = 0$.

Прежде чем решать это уравнение, обратим внимание на следующее. Искомый параметр λ определяется из квадратного уравнения. При решении квадратного уравнения корни по своему характеру могут быть: 1) действительные и различные; 2) действительные и равные; 3) мнимые.

В зависимости от характера корней получим различные ответы на поставленную геометрическую задачу. Если корни окажутся действительными и различными, то существует два различных значения параметра λ , удовлетворяющих условию задачи. Следовательно, данному пучку принадлежат две искомые прямые. Если корни окажутся действительными и равными, то существует одно значение параметра, т. е.

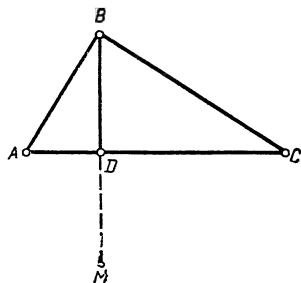


Рис. 51.

данному пучку принадлежит лишь одна искомая прямая. Если же корни уравнения окажутся мнимыми, то это означает, что действительных значений параметра λ не существует и, следовательно, не существует прямой, удовлетворяющей условиям задачи.

Уравнение $121\lambda^2 - 44\lambda + 4 = 0$ имеет единственный корень $\lambda = \frac{2}{11}$. Следовательно, существует лишь одна прямая, удовлетворяющая условиям поставленной задачи. Подставляя найденное значение λ в уравнение пучка, найдем уравнение искомой прямой, проходящей через точку пересечения данных прямых:

$$x + 2y - 19 + \frac{2}{11}(2x - y - 3) = 0 \quad \text{или} \quad 3x + 4y - 43 = 0.$$

Раздел 4

КРИВЫЕ ВТОРОГО ПОРЯДКА

§ 19. Окружность и ее уравнение

В предшествующих разделах рассматривалась прямая линия, т. е. линия, уравнение которой в декартовой прямоугольной системе координат является алгебраическим уравнением первой степени.

Будем теперь рассматривать линии, уравнения которых в декартовой системе координат являются алгебраическими уравнениями второй степени, т. е. будем рассматривать алгебраические кривые второго порядка.

Уравнение одной из таких кривых — окружности

$$(x - a)^2 + (y - b)^2 = R^2 \quad (1)$$

было выведено раньше.

Раскрывая скобки, запишем это уравнение в виде:

$$x^2 + y^2 - 2ax - 2by + a^2 + b^2 - R^2 = 0. \quad (2)$$

Это уравнение окружности является частным случаем общего уравнения второй степени:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \quad (3)$$

соответствующим $A = C \neq 0$; $B = 0$.

Обратно, если в уравнении (3) $A = C \neq 0$; $B = 0$, то после деления на A оно примет вид:

$$x^2 + y^2 + \frac{2D}{A}x + \frac{2E}{A}y + \frac{F}{A} = 0,$$

т. е.

$$\left(x + \frac{D}{A}\right)^2 + \left(y + \frac{E}{A}\right)^2 = \frac{D^2}{A^2} + \frac{E^2}{A^2} - \frac{F}{A},$$

или

$$(x - a)^2 + (y - b)^2 = p,$$

где $a = -\frac{D}{A}$; $b = -\frac{E}{A}$; $p = \frac{D^2}{A^2} + \frac{E^2}{A^2} - \frac{F}{A}$.

В случае $p \geq 0$ это будет уравнение окружности с центром (a, b) и радиусом $R = \sqrt{p}$; в случае $p = 0$ уравнению удовлетворяет единственная пара чисел (a, b) (будем говорить, что «окружность выродилась в точку»); в случае $p < 0$ уравнению не удовлетворяет ни одной пары чисел (окружность «мнимая»).

§ 20. Эллипс и его уравнение

Эллипсом называется геометрическое место точек, сумма расстояний от каждой из которых до двух данных точек F_1 и F_2 , называемых *фокусами*, есть величина постоянная¹⁾).

Для вывода уравнения эллипса выберем систему координат XOY так, чтобы фокусы эллипса F_1 и F_2 лежали на оси абсцисс, а начало координат делило бы расстояние между фокусами пополам. Обозначим $F_2F_1 = 2c$. Тогда координаты фокуса F_1 будут $c, 0$, а координаты фокуса F_2 будут $-c, 0$ (рис. 52).

Возьмем произвольную точку $M(x, y)$, лежащую на эллипсе. Соединим точку M с фокусами F_1 и F_2 . Длины отрезков MF_1 , и MF_2 обозначим соответственно через r_1 и r_2 :

$$MF_1 = r_1, \quad MF_2 = r_2.$$

Числа r_1 и r_2 называются *фокальными радиусами* точки M эллипса. Из определения эллипса следует, что сумма фокальных радиусов каждой его точки есть величина постоянная. Обозначим ее через $2a$.

В силу нашего предположения $2a > 2c$ или $a > c$. На основании определения эллипса, как геометрического места точек, можно утверждать, что для всех точек эллипса, и только для них, должно выполняться равенство:

$$r_1 + r_2 = 2a. \quad (1)$$

Определим r_1 и r_2 по формуле расстояния между двумя точками:

$$r_1 = \sqrt{(x - c)^2 + y^2}, \quad (2)$$

$$r_2 = \sqrt{(x + c)^2 + y^2}. \quad (3)$$

Подставляя найденные значения r_1 и r_2 в уравнение (1), получим:

$$\sqrt{(x - c)^2 + y^2} + \sqrt{(x + c)^2 + y^2} = 2a. \quad (4)$$

Уравнение (4) является *уравнением эллипса*. Однако полученная форма уравнения является неудобной для пользования, поэтому обычно уравнение эллипса дается в ином виде.

Пусть (x, y) — точка эллипса, т. е. равенство (4) имеет место.

Перенесем первый радикал из этого равенства направо и затем возведем обе части равенства в квадрат:

$$(x + c)^2 + y^2 = 4a^2 - 4a\sqrt{(x - c)^2 + y^2} + (x - c)^2 + y^2, \quad (5)$$

или

$$4cx = 4a^2 - 4a\sqrt{(x - c)^2 + y^2};$$

определим отсюда оставшийся радикал:

$$\sqrt{(x - c)^2 + y^2} = a - \frac{c}{a}x. \quad (6)$$

Возведем теперь обе части равенства (6) в квадрат:

$$x^2 - 2cx + c^2 + y^2 = a^2 - 2cx + \frac{c^2}{a^2}x^2$$

и умножим обе части полученного равенства на a^2 :

$$a^2x^2 + a^2c^2 + a^2y^2 = a^4 + c^2x^2.$$

¹⁾ Здесь предполагается, что эта постоянная больше F_1F_2 , так как в противном случае либо не существует точек, удовлетворяющих поставленным требованиям, либо совокупность этих точек сводится к отрезку F_1F_2 .

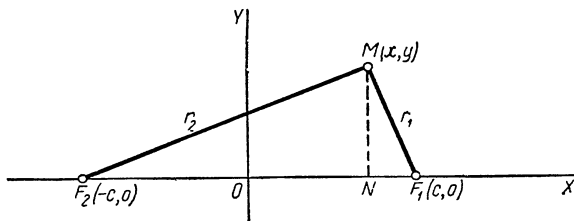


Рис. 52.

В результате преобразования получим:

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2). \quad (7)$$

Так как по условию $a > c$, то $a^2 - c^2 > 0$. Обозначим разность $a^2 - c^2$ (как величину положительную) через b^2 . Позже мы раскроем геометрический смысл положительного числа b , определяемого равенством

$$b^2 = a^2 - c^2. \quad (8)$$

Очевидно, что

$$b^2 < a^2.$$

Вводя в равенство (7) $b^2 = a^2 - c^2$ и разделив все на a^2b^2 , получим:

$$\begin{aligned} b^2x^2 + a^2y^2 &= a^2b^2, \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} &= 1. \end{aligned}$$

Пусть теперь x и y — любые действительные числа.

Рассмотрим уравнение

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (9)$$

По доказанному, всякая пара чисел x, y , удовлетворяющая уравнению (4), удовлетворяет и уравнению (9). Можно показать, что и, наоборот, всякая пара чисел x, y , удовлетворяющая уравнению (9), удовлетворяет уравнению (4) (для этого достаточно произвести сделанные выше выкладки в обратном порядке). Таким образом, уравнению (9) удовлетворяют координаты точек данного эллипса, и только они, т. е. уравнение (9) является уравнением этого эллипса.

Уравнение (9) называется *каноническим уравнением эллипса*.

§ 21. Исследование уравнения эллипса

Получив каноническое уравнение эллипса, приступим к изучению формы эллипса. В уравнение эллипса обе координаты входят только в четных степенях, поэтому, если пара чисел (x_0, y_0) удовлетворяет уравнению (9), то пары $(x_0, -y_0)$, $(-x_0, y_0)$, $(-x_0, -y_0)$ также удовлетворяют этому уравнению. Другими словами, если точка $M_0(x_0, y_0)$ лежит на эллипсе, то точки $M_1(x_0, -y_0)$, $M_2(-x_0, y_0)$, $M_3(-x_0, -y_0)$, симметричные точке M_0 относительно оси OX , оси OY и начала O , также лежат на эллипсе. Отсюда следует, что эллипс является кривой, симметричной относительно обеих осей координат и начала координат. Это позволяет изучение формы и построение эллипса ограничить первым квадрантом, а затем получившуюся кривую с помощью зеркального ото-

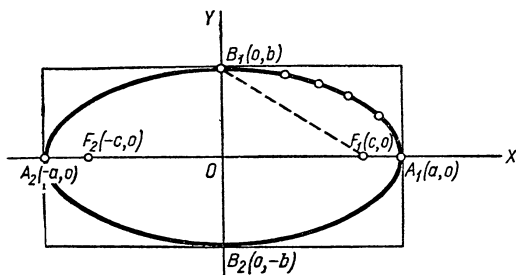


Рис. 53.

бражения построить во всех четырех квадрантах. В случае канонического задания эллипса [(9), § 20] координатные оси являются *осями симметрии* эллипса. Точка пересечения осей симметрии называется *центром* эллипса.

Построим точки $A_1(a, 0)$, $A_2(-a, 0)$, $B_1(0, b)$, $B_2(0, -b)$, а затем через точки A_1, A_2, B_1 и B_2 проведем прямые, параллельные координатным осям, до их взаимного пересечения и таким образом построим прямоугольник (рис. 53). Затем раствором циркуля, равным a , из точки B_1

как из центра проведем засечки на оси абсцисс. Нетрудно убедиться, что полученные точки F_1 и F_2 являются фокусами эллипса.

$B_1F_1=a$, $OB_1=b$, следовательно, $a^2-OF_1^2=b^2$ и на основании равенства (8) (§ 21) $a^2-c^2=b^2$ имеем $OF_1=c$.

Из канонического уравнения эллипса $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ выразим y через x :

$$y=\pm\frac{b}{a}\sqrt{a^2-x^2}. \quad (1)$$

Так как изучение формы эллипса достаточно провести в первом квадранте, то в этом равенстве надо взять лишь знак плюс, т. е. $y=\frac{b}{a}\sqrt{a^2-x^2}$, и предполагать $x\geq 0$.

1) При $x=0$ имеем $y=b$. Следовательно, точка $B_1(0, b)$ лежит на эллипсе.

2) При возрастании x от 0 до a y убывает.

3) При $x=a$ имеем $y=0$. Следовательно, точка $A_1(a, 0)$ лежит на эллипсе.

4) При $x>a$ y получает мнимые значения. Следовательно, точек эллипса, у которых $x>a$, не существует.

Дадим переменной x несколько значений, $0<x<a$, и, получив соответствующие значения y , $b>y>0$, построим ряд точек, принадлежащих эллипсу. Учитывая высказанные выше соображения и соединяя найденные точки эллипса плавной линией, получим дугу эллипса B_1A_1 в первом квадранте.

Произведя зеркальные отражения дуги B_1A_1 относительно координатных осей, получим весь эллипс.

Отсюда следует, что эллипс представляет собой замкнутую кривую, не выходящую из пределов начерченного прямоугольника, с двумя взаимно перпендикулярными осями симметрии.

Отрезок A_2A_1 и его длина $2a$ называются *большой осью* эллипса, отрезок OA_1 и его длина a называются *большой полуосью* эллипса. Отрезок B_2B_1 и его длина $2b$ называются *малой осью* эллипса; отрезок OB_1 и его длина b называются *малой полуосью* эллипса. Длина отрезка F_2P_1 , т. е. число $2c$, называется *фокусным расстоянием*. Точки пересечения эллипса с его осями A_1, A_2, B_1, B_2 называются *вершинами* эллипса, а точка пересечения его осей (как уже упоминалось) называется *центром* эллипса¹⁾.

Примечание. Если $a=b$, то уравнение эллипса имеет вид $\frac{x^2}{a^2}+\frac{y^2}{a^2}=1$ или $x^2+y^2=a^2$. Это — уравнение окружности. Можно сказать, что окружность является частным случаем эллипса.

Сравнивая уравнение окружности радиуса a с центром в начале координат $x^2+y^2=a^2$ или $y=\pm\sqrt{a^2-x^2}$ с уравнением (1), заключаем, что ордината точки эллипса и ордината точки окружности с той же абсциссой в той же полуплоскости (верхней или нижней) связаны зависимостью $y_{эл.}=\frac{b}{a}y_{окр.}$. Это показывает, что эллипс можно рассматривать как «сплюснутую» окружность.

¹⁾ Вообще *центром* какой-нибудь линии называется точка M_0 , обладающая следующим свойством: если точка M принадлежит линии, то точка, симметричная M относительно M_0 , также принадлежит этой линии. Линии, имеющие центр, называются *центральными*, линии, не имеющие центра, — *нецентральными*. Центр линии, если он есть, не обязательно единственный. Например, каждая точка прямой линии является ее центром.

§ 22. Эксцентриситет и фокальные радиусы эллипса

Эксцентриситетом эллипса называется отношение фокусного расстояния к длине большой оси эллипса. Эксцентриситет обычно обозначается буквой ϵ :

$$\frac{2c}{2a} = \epsilon \quad \text{или} \quad \frac{c}{a} = \epsilon. \quad (1)$$

Так как у эллипса $c < a$, то эксцентриситет эллипса меньше единицы. Определяя c из уравнения эллипса, получим для ϵ^2 следующее выражение:

$$\epsilon^2 = \frac{c^2}{a^2} = \frac{a^2 - b^2}{a^2} = 1 - \left(\frac{b}{a}\right)^2. \quad (2)$$

Из этого равенства видно, что чем меньше отношение $\frac{b}{a}$, т. е. чем больше эксцентриситет, тем эллипс становится более «вытянутым». Наоборот, чем больше отношение $\frac{b}{a}$, тем меньше эксцентриситет и эллипс является менее «вытянутым» (менее «сжатым»). Таким образом, эксцентриситет характеризует степень сжатия эллипса. В предельном случае, когда $b = a$, т. е. когда эллипс обращается в окружность, его эксцентриситет обращается в нуль.

Фокальные радиусы. Из равенства (6) § 20 имеем:

$$r_1 = \sqrt{(x - c)^2 + y^2} = a - \frac{c}{a}x,$$

но $r_1 + r_2 = 2a$, следовательно, $r_2 = a + \frac{c}{a}x$. Таким образом,

$$\begin{aligned} r_1 &= a - \frac{c}{a}x, \\ r_2 &= a + \frac{c}{a}x. \end{aligned} \quad (3)$$

Равенства (3) позволяют выражать фокальные радиусы точки $M(x, y)$ через x (без радикалов).

Так как $\epsilon = \frac{c}{a}$, то формулы (3) можно переписать так:

$$\begin{aligned} r_1 &= a - \epsilon x, \\ r_2 &= a + \epsilon x. \end{aligned} \quad (4)$$

§ 23. Гипербола и ее уравнение

Гиперболой называется геометрическое место точек, разность расстояний от каждой из которых до двух данных, точек F_1 и F_2 , называемых *фокусами*, есть величина постоянная¹⁾.

Для вывода уравнения гиперболы возьмем систему координат $ХОУ$ так, чтобы фокусы гиперболы F_1 и F_2 лежали на оси абсцисс, а начало координат делило отрезок F_2F_1 ($F_1F_2 = 2c$) пополам (рис. 54). Тогда координаты фокуса F_1 будут $(c, 0)$, а фокуса F_2 — $(-c, 0)$.

Возьмем точку $M(x, y)$, лежащую на гиперболе, и проведем отрезки MF_1 и MF_2 . Длину отрезка MF_1 обозначим через r_1 , а длину отрезка MF_2 — через r_2 :

$$MF_1 = r_1, \quad MF_2 = r_2.$$

¹⁾ Здесь предполагается, что эта постоянная меньше F_1F_2 , так как в противном случае либо не существовало бы точек, удовлетворяющих требованиям, либо совокупность этих точек сводилась бы к двум лучам прямой, проходящей через F_1 и F_2 .

Числа r_1 и r_2 называются *фокальными радиусами* точки M гиперболы. По определению гиперболы, для любой ее точки разность фокальных радиусов есть величина постоянная (обозначим ее через $2a$). В силу нашего предположения $2a < 2c$ или $a < c$.

На основании определения гиперболы как геометрического места точек, можно утверждать, что для всех точек гиперболы, и только для них, должно выполняться равенство:

$$r_2 - r_1 = \pm 2a. \quad (1)$$

По формуле расстояния между двумя точками определим:

$$MF_1 = r_1 \text{ и } MF_2 = r_2:$$

$$r_1 = \sqrt{(x-c)^2 + y^2},$$

$$r_2 = \sqrt{(x+c)^2 + y^2}.$$

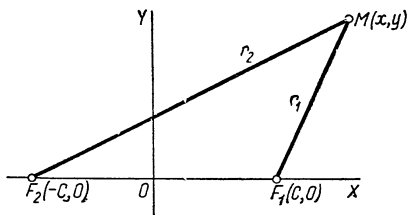


Рис. 54.

Подставляя найденные значения r_1 и r_2 в уравнение (1), получим:

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = \pm 2a. \quad (2)$$

Уравнение (2) является *уравнением гиперболы*. Так же как и уравнение эллипса, приведем уравнение (2) к более удобному виду, для чего уединим первый радикал и затем возведем обе части равенства (2) в квадрат. При этом

$$(x+c)^2 + y^2 = 4a^2 \pm 4a\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2,$$

т. е.

$$4cx = 4a^2 \pm 4a\sqrt{(x-c)^2 + y^2}.$$

Отсюда

$$\pm \sqrt{(x-c)^2 + y^2} = \frac{c}{a}x - a. \quad (3)$$

Возведем обе части равенства (3) в квадрат:

$$x^2 - 2cx + c^2 + y^2 = \frac{c^2}{a^2}x^2 - 2cx + a^2$$

и освободимся от знаменателя. Получим:

$$a^2x^2 + a^2c^2 + a^2y^2 = c^2x^2 + a^4,$$

т. е.

$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2). \quad (4)$$

Так как по условию $a < c$, то $c^2 - a^2 > 0$. Обозначим $c^2 - a^2$, как величину положительную, через b^2 . При исследовании уравнения гиперболы будет раскрыта геометрическая сущность положительного параметра b , определяемого равенством

$$c^2 - a^2 = b^2. \quad (5)$$

Подставив в равенство (4) $b^2 = c^2 - a^2$, а затем деля все его члены на a^2b^2 , получим:

$$b^2x^2 - a^2y^2 = a^2b^2,$$

т. е.

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (6)$$

Уравнению (6) будут удовлетворять координаты каждой точки, лежащей на гиперболы. Можно показать, что координаты точек, не при-

надлежащих гиперболе, уравнению (6) не удовлетворяют. Следовательно, уравнение (6) является уравнением рассматриваемой гиперболы. Уравнение (6) называется *каноническим уравнением гиперболы*.

§ 24. Исследование уравнения гиперболы

Получив каноническое уравнение гиперболы, приступим к изучению формы гиперболы. Прежде всего заметим, что в уравнение гиперболы обе координаты входят только в четных степенях. Следовательно, если некоторая точка $M(x_0, y_0)$ лежит на гиперболе, то на гиперболе будут также лежать точки $M_1(x_0, -y_0)$; $M_2(-x_0, y_0)$; $M_3(-x_0, -y_0)$. Отсюда следует, что гипербола является кривой, симметричной относительно обеих координатных осей и начала координат. Это позволяет изучение формы гиперболы ограничить первым квадрантом, а затем получившуюся кривую с помощью зеркального изображения построить во всех четырех квадрантах. В случае канонического задания гиперболы координатные оси являются

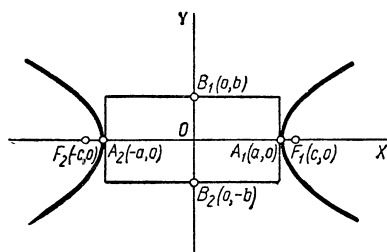


Рис. 55.

осями симметрии гиперболы, а начало координат центром симметрии. Точка пересечения осей симметрии называется *центром* гиперболы. Таким образом, гипербола, как и эллипс, — центральная кривая.

От начала координат на оси абсцисс вправо и влево отложим отрезок, длина которого равна a , и построим точки $A_1(a, 0)$ и $A_2(-a, 0)$, а на оси ординат вверх и вниз отложим отрезок длины b и построим точки $B_1(0, b)$ и

$B_2(0, -b)$. Затем через точки A_1, A_2, B_1, B_2 проведем прямые, параллельные координатным осям, до их взаимного пересечения и таким образом построим прямоугольник (рис. 55), который называют *основным прямоугольником* гиперболы.

Раствором циркуля, равным расстоянию A_1B_1 , из начала координат, как из центра, сделаем засечки на оси абсцисс. При этом мы найдем точки F_1 и F_2 . Действительно, из прямоугольного треугольника OA_1B_1 : $OA_1 = a$, $OB_1 = b$. Следовательно, на основании равенства (5), § 23:

$$a^2 + b^2 = c^2, \text{ т. е. } B_1A_1 = c.$$

Определим теперь y из канонического уравнения гиперболы

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1:$$

$$y = \pm \frac{b}{a} \sqrt{x^2 - a^2}. \quad (1)$$

Так как исследование гиперболы будет вестись в первом квадранте, то в этом равенстве надо перед корнем взять знак плюс:

$$y = \frac{b}{a} \sqrt{x^2 - a^2} \quad (2)$$

и рассматривать $x \geq 0$.

1) При $0 \leq x < a$ y получает мнимые значения, следовательно, точек гиперболы с абсциссами x , $0 \leq x < a$, не существует.

2) Если $x = a$, то $y = 0$. Следовательно, точка $A_1(a, 0)$ принадлежит гиперболе.

3) Если $x > a$, то $y > 0$, причем при возрастании x возрастает и y . Когда x неограниченно возрастает, y также неограниченно возрастает. Следовательно, при неограниченном возрастании x ветвь гиперболы уходит в бесконечность.

Отсюда следует, что гипербола представляет собой кривую, состоящую из двух бесконечных ветвей¹⁾ с двумя взаимно перпендикулярными осями симметрии, причем ни одной точки гиперболы не находится внутри основного прямоугольника.

Отрезок A_2A_1 и его длина $2a$ называются *действительной осью* гиперболы, отрезок OA_1 и его длина a называются *действительной полуосью* гиперболы. Отрезок B_2B_1 и его длина $2b$ называются *мнимой осью* гиперболы; отрезок OB_1 и его длина b называются *мнимой полуосью* гиперболы. Длина $2c$ отрезка F_2F_1 называется *фокусным расстоянием*. Точки пересечения гиперболы с действительной осью A_1 и A_2 называются *вершинами* гиперболы.

Асимптоты гиперболы. Возьмем уравнение гиперболы в канонической форме:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (3)$$

и построим для нее основной прямоугольник (рис. 56). *Асимптотами*²⁾ гиперболы называются прямые, имеющие уравнения

$$y = \frac{b}{a}x \quad \text{и} \quad y = -\frac{b}{a}x. \quad (4)$$

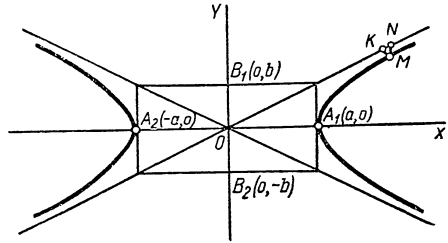


Рис. 56.

Эти прямые являются диагоналями основного прямоугольника. Построим гиперболу $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ и рассмотрим какую-нибудь точку $M(x, y)$, лежащую на гиперболе в первом квадранте.

Выясним, как в первом квадранте по мере возрастания x будет изменяться расстояние от точки M гиперболы до асимптоты $y = \frac{b}{a}x$. Обозначим через N точку асимптоты с абсциссой x : $N(x, Y)$, где $Y = \frac{b}{a}x$. Тогда

$$Y - y = \frac{b}{a}x - \frac{b}{a}\sqrt{x^2 - a^2} = \frac{b}{a}(x - \sqrt{x^2 - a^2}). \quad (5)$$

Так как $a \leq x$, то в скобках первое слагаемое всегда больше второго, следовательно, $Y - y > 0$, а это означает, что при одной и той же абсциссе точка гиперболы лежит под соответствующей точкой асимптоты.

Преобразуя равенство (5):

$$MN = Y - y = \frac{b}{a}(x - \sqrt{x^2 - a^2}) = \frac{b}{a} \frac{a^2}{x + \sqrt{x^2 + a^2}} = \frac{ab}{x + \sqrt{x^2 + a^2}}, \quad (6)$$

убеждаемся, что длина отрезка MN по мере возрастания x уменьшается, и когда x неограниченно растет, то MN стремится к нулю.

Так как MN больше расстояния MK от точки M до асимптоты, то при этом MK и подавно стремится к нулю.

Аналогичное рассуждение можно провести в любом квадранте. Отметим общее определение асимптоты какой-нибудь линии. Прямая L называется *асимптотой* линии C , если расстояние от точки M линии C до этой прямой стремится к нулю по мере неограниченного удаления точки M от начала координат.

Итак, прямые $y = \pm \frac{b}{a}x$ в смысле данного определения являются асимптотами гиперболы

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (7)$$

¹⁾ Для правой ветви $r_2 - r_1 = 2a$, для левой $r_2 - r_1 = -2a$.

²⁾ Общее определение асимптот кривых приведено на стр. 243.

При построении гиперболы обычно строят основной прямоугольник и проводят асимптоты, так как они позволяют точнее вычертить гиперболу.

Равнобочная гипербола. Возьмем каноническое уравнение гиперболы

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

В случае, когда $a = b$, уравнение гиперболы имеет вид $\frac{x^2}{a^2} - \frac{y^2}{a^2} = 1$ или

$$x^2 - y^2 = a^2. \quad (8)$$

Гипербола, у которой полуоси a и b равны, называется *равнобочной* гиперболой. Уравнение (8) называется уравнением равнобочной гиперболы.

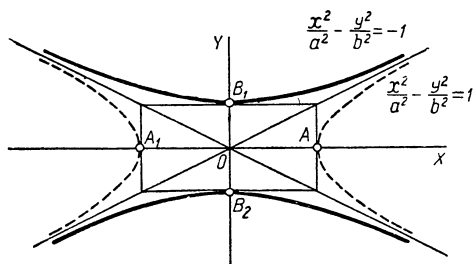


Рис. 57.

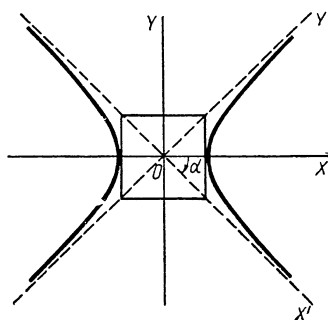


Рис. 58.

Так как основным прямоугольником этой гиперболы является квадратом, то асимптоты равнобочной гиперболы будут перпендикулярны друг другу.

Сопряженная гипербола. Рассмотрим уравнение

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1. \quad (9)$$

Представим уравнение (9) в следующем виде:

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1. \quad (10)$$

Очевидно, что уравнение (10) представляет собой уравнение гиперболы, у которой действительной осью является ось ординат, а мнимой — ось абсцисс.

Построим основной прямоугольник, проведем асимптоты и построим гиперболу (9) (рис. 57). Далее в той же системе координат построим (пунктиром) гиперболу

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Очевидно, что гиперболы $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ и $\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$ имеют общие асимптоты. Такие гиперболы называются *сопряженными*.

Уравнение равнобочной гиперболы, асимптотами которой служат оси координат. Возьмем уравнение равнобочной гиперболы $x^2 - y^2 = a^2$ и рассмотрим уравнение этой гиперболы в новой системе координат $X'OY'$, полученной из старой поворотом осей координат на угол $\alpha = -\frac{\pi}{4}$ (рис. 58).

Используем для этого формулы поворота осей координат

$$x = x' \cos \alpha - y' \sin \alpha, \quad y = x' \sin \alpha + y' \cos \alpha.$$

Подставим значения x, y в уравнение гиперболы $x^2 - y^2 = a^2$:

$$\begin{aligned} \left[x' \cos \left(-\frac{\pi}{4} \right) - y' \sin \left(-\frac{\pi}{4} \right) \right]^2 - \left[x' \sin \left(-\frac{\pi}{4} \right) + y' \cos \left(-\frac{\pi}{4} \right) \right]^2 &= a^2, \\ \frac{1}{2} (x' + y')^2 - \frac{1}{2} (-x' + y')^2 &= a^2, \\ x'y' &= \frac{a^2}{2}. \end{aligned} \quad (11)$$

Обозначая $\frac{a^2}{2} = c$, получим $x'y' = c$.

Уравнение равнобочной гиперболы, для которой координатные оси OX и OY являются асимптотами, будет иметь вид:

$$xy = c$$

или

$$y = \frac{c}{x}. \quad (12)$$

§ 25. Эксцентриситет и фокальные радиусы гиперболы

Эксцентриситетом гиперболы называется отношение фокусного расстояния к длине ее действительной оси:

$$\frac{2c}{2a} = \varepsilon \quad \text{или} \quad \frac{c}{a} = \varepsilon.$$

Так как у гиперболы $c > a$, то эксцентриситет гиперболы больше единицы. Так же как и у эллипса, эксцентриситет характеризует отношение сторон основного прямоугольника, а следовательно, и форму самой гиперболы.

Фокальные радиусы. Из равенства (3) § 23 имеем (для правой ветви)

$$r_1 = \sqrt{(x-c)^2 + y^2} = \frac{c}{a} x - a,$$

но

$$r_2 - r_1 = 2a.$$

Следовательно,

$$r_2 = \frac{c}{a} x + a.$$

Таким образом, получаем формулы, выражающие фокальные радиусы любой точки $M(x, y)$ правой ветви гиперболы через x :

$$\left. \begin{aligned} r_1 &= \frac{c}{a} x - a, \\ r_2 &= \frac{c}{a} x + a. \end{aligned} \right\} \quad (1)$$

Для левой ветви эти формулы примут вид:

$$\left. \begin{aligned} r_1 &= -\left(\frac{c}{a} x - a \right), \\ r_2 &= -\left(\frac{c}{a} x + a \right). \end{aligned} \right\} \quad (2)$$

Выражая формулы (1) и (2) через эксцентриситет, получим для точек правой ветви гиперболы

$$\left. \begin{aligned} r_1 &= \varepsilon x - a, \\ r_2 &= \varepsilon x + a, \end{aligned} \right\} \quad (3)$$

для точек левой ветви гиперболы:

$$\left. \begin{aligned} r_1 &= -(\varepsilon x - a), \\ r_2 &= -(\varepsilon x + a). \end{aligned} \right\} \quad (4)$$

§ 26. Парабола и ее уравнение

Параболой называется геометрическое место точек, каждая из которых находится на одинаковом расстоянии от данной точки, называемой *фокусом*, и данной прямой, называемой *директрисой*.

Для вывода уравнения параболы за ось OX возьмем прямую, проходящую через фокус перпендикулярно директрисе. За положительное направление оси абсцисс возьмем направление от директрисы к фокусу.

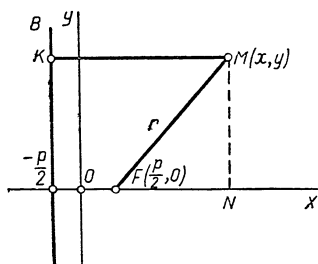


Рис. 59.

За начало координат возьмем точку O , которая делит пополам отрезок от директрисы до фокуса. Длину этого отрезка, который называется *параметром параболы*, обозначим через p . Фокус F будет иметь координаты $\frac{p}{2}, 0$, а координаты точки оси OX , через

которую проходит директриса, будут $-\frac{p}{2}, 0$.

Возьмем произвольную точку $M(x, y)$, лежащую на параболе, соединим ее прямой с точкой F , а затем опустим из точки M на директрису перпендикуляр MK . Длина отрезка, соединяющего точку $M(x, y)$ параболы с фокусом, называется *фокальным радиусом* этой точки и обозначается через r (рис. 59).

Согласно определению параболы

$$FM = KM. \quad (1)$$

Определяя FM и KM по формуле расстояния между двумя точками, получим

$$FM = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2}, \quad KM = \sqrt{\left(x + \frac{p}{2}\right)^2 + (y - 0)^2}.$$

Следовательно,

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \sqrt{\left(x + \frac{p}{2}\right)^2}. \quad (2)$$

Уравнению (2) будут удовлетворять координаты каждой точки параболы.

Приведем уравнение параболы к более удобному виду, для чего возведем обе части равенства (2) в квадрат:

$$x^2 - px + \frac{p^2}{4} + y^2 = \frac{p^2}{4} + px + x^2.$$

Откуда

$$y^2 = 2px. \quad (3)$$

Уравнение (3) называется *каноническим уравнением параболы*¹⁾. Составляя равенства (1) и (2), можно выразить фокальный радиус точки $M(x, y)$ параболы через абсциссу этой точки:

$$r = \frac{p}{2} + x. \quad (4)$$

§ 27. Исследование уравнения параболы

Для определения вида параболы определим y из канонического уравнения параболы:

$$y = \pm \sqrt{2px}.$$

¹⁾ Можно показать, что ему удовлетворяют только координаты точек, лежащих на данной параболе.

Из уравнения (3) § 26 следует, что x не может быть отрицательным. При $x=0$ имеем $y=0$, следовательно, точка $O(0, 0)$ лежит на параболе. Затем заключаем, что каждому значению $x > 0$ соответствует два значения y , равных по абсолютной величине, но противоположных по знаку. Следовательно, парабола представляет собой кривую, расположенную вправо от начала координат и симметричную относительно оси абсцисс (рис. 60). Из формулы (3) § 26 следует, что по мере возрастания x возрастает и $|y|$, и когда x неограниченно растет, то и y по абсолютной величине неограниченно растет.

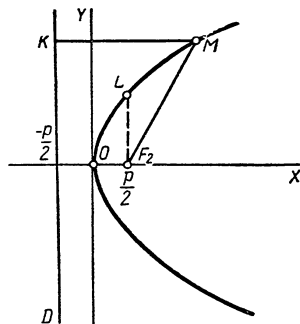


Рис. 60.

У параболы, заданной каноническим уравнением $y^2 = 2px$, осью симметрии является ось абсцисс. Точка пересечения параболы с осью симметрии называется *вершиной* параболы. В данном случае вершина параболы лежит в начале координат. Заметим, что у параболы одна вершина, у гиперболы — две, у эллипса — четыре. Проведем на рис. 60 фокальный радиус перпендикулярно оси симметрии и определим

его длину LF_2 по формуле (4) § 26. Так как абсцисса точки L равна $\frac{p}{2}$, то $r = p$. Следовательно, число p равняется длине фокального радиуса, перпендикулярного к оси симметрии. В связи с этим число p называют *фокальным параметром* параболы.

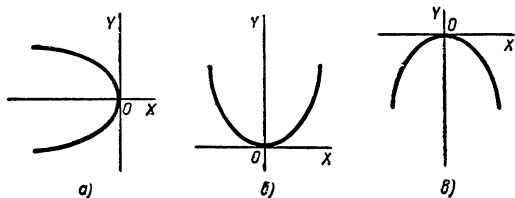


Рис. 61.

Парабола, уравнение которой $y^2 = 2px$, $p > 0$, является кривой, расположенной справа от оси ординат.

Кривая, уравнение которой $y^2 = -2px$, $p > 0$, будет также параболой. Вершина этой параболы лежит в начале координат, осью симметрии

является ось абсцисс. Все точки этой параболы лежат слева от оси ординат (рис. 61, а). Рассуждая аналогичным образом, заключаем, что уравнение $x^2 = 2py$, $p > 0$, является уравнением параболы, вершина которой лежит в начале координат, осью симметрии является ось ординат (рис. 61, б). Эта парабола лежит выше оси абсцисс. Уравнение же вида $x^2 = -2py$, $p > 0$, является уравнением параболы, лежащей ниже оси абсцисс, с вершиной в начале координат. Ось симметрии этой параболы является ось ординат (рис. 61, в).

Примечание. Условимся, наглядности ради, говорить, что «ветви» параболы $y^2 = 2px$ ($p > 0$) «направлены вправо», «ветви» параболы $x^2 = 2py$ ($p > 0$) «направлены вверх» и т. д.

§ 28. Общее свойство кривых второго порядка — эллипса, гиперболы и параболы

Директрисы эллипса и гиперболы. Построим эллипс, заданный каноническим уравнением $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Затем построим две прямые, перпендикулярные к большой оси эллипса и расположенные симметрично относительно центра на расстоянии $\frac{a}{e}$ (рис. 62). Эти прямые, уравнения которых будут: $x = \frac{a}{e}$ и $x = -\frac{a}{e}$, называются *директрисами эллипса*.

При их построении следует учесть, что $\frac{a}{\varepsilon} > a$, так как эксцентриситет эллипса $\varepsilon < 1$.

Правая директриса, уравнение которой $x = \frac{a}{\varepsilon}$, будет проходить правее вершины эллипса A_1 , а левая директриса, уравнение которой $x = -\frac{a}{\varepsilon}$, —

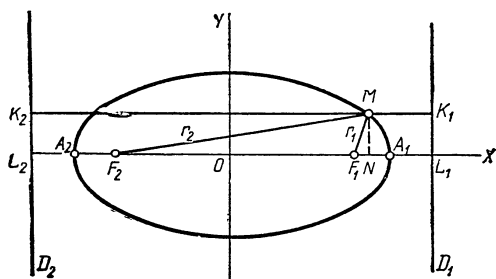


Рис. 62.

левее вершины эллипса A_2 (рис. 62).

Построим гиперболу, заданную каноническим уравнением:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

и две прямые, перпендикулярные действительной оси гиперболы и симметрично расположенные относительно центра на расстоянии, равном $\frac{a}{\varepsilon}$ (рис. 63).

Эти прямые (их уравнения $x = \frac{a}{\varepsilon}$ и $x = -\frac{a}{\varepsilon}$) называются *директрисами гиперболы* (соответственно, *правой* и *левой*). При их построении следует учесть, что $\frac{a}{\varepsilon} < a$, так как эксцентриситет гиперболы $\varepsilon > 1$.

Правая директриса гиперболы ($x = \frac{a}{\varepsilon}$) будет проходить левее правой вершины гиперболы A_1 , а левая директриса гиперболы ($x = -\frac{a}{\varepsilon}$) будет проходить правее левой вершины гиперболы A_2 (рис. 63).

С помощью директрис и эксцентриситета можно выявить общее свойство, присущее кривым второго порядка: эллипсу, гиперболе и параболе. Имеет место следующая теорема: отношение расстояний от произвольной точки $M(x, y)$ любой из этих кривых до фокуса и до соответствующей директрисы есть величина постоянная, равная эксцентриситету кривой. Докажем эту теорему последовательно для эллипса, гиперболы и параболы.

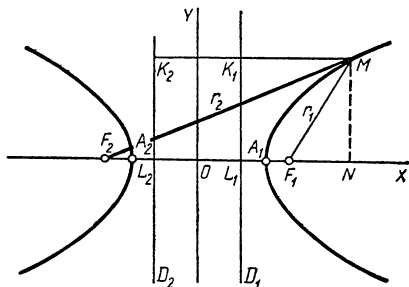


Рис. 63.

Доказательство. Пусть у эллипса F_1 — правый фокус, прямая D_1L_1 — правая директриса, F_2 — левый фокус, D_2L_2 — левая директриса (рис. 62). Возьмем на эллипсе произвольную точку $M(x, y)$, соединим ее отрезками MF_1 и MF_2 ($MF_1 = r_1$, $MF_2 = r_2$) с фокусами и опустим из нее перпендикуляры MK_1 и MK_2 на обе директрисы ($MK_1 = d_1$ и $MK_2 = d_2$) и на ось OX . Требуется доказать, что

$$\frac{r_1}{d_1} = \frac{r_2}{d_2} = \varepsilon. \quad (1)$$

На основании выведенных ранее формул имеем:

$$r_1 = a - \varepsilon x, \quad r_2 = a + \varepsilon x,$$

$$d_1 = \overline{OL_1} - \overline{ON} = \frac{a}{\varepsilon} - x^1),$$

$$d_2 = \overline{L_2O} + \overline{ON} = \frac{a}{\varepsilon} + x. \quad (2)$$

¹⁾ Здесь N — основание перпендикуляра, опущенного из точки M на ось OX .

Вычисляя отношения $\frac{r_1}{d_1}$ и $\frac{r_2}{d_2}$, получим:

$$\frac{r_1}{d_1} = \frac{a - \epsilon x}{\frac{a}{\epsilon} - x} = \frac{(a - \epsilon x) \epsilon}{a - \epsilon x} = \epsilon; \quad \frac{r_2}{d_2} = \frac{a + \epsilon x}{\frac{a}{\epsilon} + x} = \frac{(a + \epsilon x) \epsilon}{a + \epsilon x} = \epsilon.$$

Таким образом, данная теорема для эллипса доказана.

Пусть у гиперболы F_1 — правый фокус, D_1L_1 — соответствующая ему правая директриса, F_2 — левый фокус, D_2L_2 — соответствующая ему левая директриса (рис. 63). Возьмем на гиперболе произвольную точку $M(x, y)$ и соединим ее с фокусами: $F_1M = r_1$, $F_2M = r_2$, затем из точки M опустим перпендикуляры на обе директрисы: $K_1M = d_1$ и $K_2M = d_2$. Требуется доказать, что

$$\frac{r_1}{d_1} = \frac{r_2}{d_2} = \epsilon. \quad (1)$$

Применяя выведенные выше формулы, получаем:

$$r_1 = \epsilon x - a, \quad r_2 = \epsilon x + a,$$

$$d_1 = \overline{ON} - \overline{OL_1} = x - \frac{a}{\epsilon},$$

$$d_2 = \overline{ON} + \overline{L_2O} = \frac{a}{\epsilon} + x$$

и, следовательно,

$$\frac{r_1}{d_1} = \frac{\epsilon x - a}{x - \frac{a}{\epsilon}} = \frac{(\epsilon x - a) \epsilon}{\epsilon x - a} = \epsilon, \quad \frac{r_2}{d_2} = \frac{\epsilon x + a}{x + \frac{a}{\epsilon}} = \frac{(\epsilon x + a) \epsilon}{\epsilon x + a} = \epsilon.$$

Таким образом, теорема доказана и для гиперболы.

Что касается параболы, являющейся геометрическим местом таких точек, которые равно удалены от фокуса и директрисы, то для любой точки параболы будет справедливо равенство $\frac{r}{d} = 1$, где d расстояние от точки параболы до директрисы. Иначе, отношение расстояний от любой точки параболы до фокуса и до директрисы равно единице. По аналогии с остальными кривыми второго порядка, это постоянное отношение называют *эксцентриситетом параболы*. Следовательно, эксцентриситет параболы равен единице.

Таким образом, сформулированная теорема доказана для эллипса, гиперболы и параболы. Как будет показано в § 39 и 40, любая кривая второго порядка (если не считать случаев вырождения) является эллипсом, гиперболой или параболой. В связи с этим для рассмотренных выше кривых второго порядка можно дать следующее общее определение: кривые второго порядка суть геометрические места точек, отношение расстояний которых до фокуса и до соответствующей директрисы есть величина постоянная, равная ϵ , причем у эллипса $\epsilon < 1$, у гиперболы $\epsilon > 1$, у параболы $\epsilon = 1$.

Полярное уравнение кривой второго порядка. Пользуясь общим свойством эллипсов, гипербол и парабол, выведем общее уравнение этих кривых второго порядка в полярных координатах при некотором специальном выборе полярной системы координат.

Пусть дана какая-нибудь из указанных линий (эллипс, ветвь гиперболы или парабола). Возьмем фокус F кривой (любой, если их два) и соответствующую ему директрису L . (Если рассматривается ветвь гиперболы, то берется фокус и директриса, ближайшие к этой ветви.)

Введем полярную систему координат так, чтобы полюс O совпал с фокусом F , а полярная ось была направлена по оси симметрии кривой в сторону, противоположную директрисе L .

Возьмем на кривой произвольную точку $M(\rho, \varphi)$, соединим ее отрезком FM с фокусом и опустим перпендикуляр MK на директрису. Кроме того, из точки F проведем перпендикуляр FR к полярной оси до пересечения с кривой в точке R , а из точки R опустим перпендикуляр RQ на директрису (рис. 64).

Обозначим $FR = p$ и будем называть это число *фокальным параметром*.

На основании общего свойства кривых второго порядка

$$\frac{FM}{KM} = \varepsilon; \quad FM = \rho; \quad KM = DF + \rho \cos \varphi.$$

По тем же соображениям

$$\frac{FR}{QR} = \varepsilon \quad \text{или} \quad \frac{p}{DF} = \varepsilon, \quad \text{откуда} \quad DF = \frac{p}{\varepsilon}.$$

Подставляя найденные выражения для FM и KM в равенство $\frac{FM}{KM} = \varepsilon$, получим:

$$\frac{\rho}{\frac{p}{\varepsilon} + \rho \cos \varphi} = \varepsilon, \quad \rho = \frac{p}{1 - \varepsilon \cos \varphi}. \quad (3)$$

Уравнение (3) называется *уравнением кривой второго порядка в полярных координатах*. При $\varepsilon < 1$ кривая является эллипсом, при $\varepsilon > 1$ — ветвью гиперболы, при $\varepsilon = 1$ — параболой.

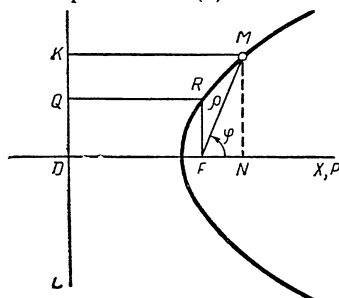


Рис. 64.

Фокальный параметр p из уравнения параболы определяется непосредственно. Для того чтобы фокальный параметр выразить через параметры эллипса или гиперболы, следует заметить, что фокальный параметр p является ординатой точки кривой, абсцисса которой равна абсциссе соответствующего фокуса (в выбранной при выводе канонического уравнения соответствующей кривой системе XOY).

Подставляя вместо координат точки $M(x, y)$ в уравнение эллипса $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ координаты точки $(-c, p)$, получим:

$$\frac{c^2}{a^2} + \frac{p^2}{b^2} = 1 \quad \text{или} \quad \frac{p^2}{b^2} = \frac{a^2 - c^2}{a^2} = \frac{b^2}{a^2},$$

откуда

$$p = \frac{b^2}{a}. \quad (4)$$

Аналогично, подставляя в уравнение гиперболы координаты точки (c, p) , получим:

$$\frac{c^2}{a^2} - \frac{p^2}{b^2} = 1 \quad \text{или} \quad \frac{p^2}{b^2} = \frac{c^2 - a^2}{a^2} = \frac{b^2}{a^2},$$

откуда следует соотношение

$$p = \frac{b^2}{a}. \quad (5)$$

В заключение рассмотрим несколько задач на кривые второго порядка.

Задача 1. Дано уравнение гиперболы $9x^2 - 16y^2 = 144$. Найти длины ее осей, координаты фокусов, эксцентриситет; составить уравнения директрис и асимптот гиперболы.

Решение. Приведем уравнение гиперболы к каноническому виду и определим как параметры гиперболы, так и расстояние c от начала координат до фокуса:

$$\frac{9x^2}{144} - \frac{16y^2}{144} = 1 \quad \text{или} \quad \frac{x^2}{16} - \frac{y^2}{9} = 1,$$

откуда $a=4$, $b=3$, $c=\sqrt{a^2+b^2}=5$, эксцентриситет $\varepsilon = \frac{c}{a} = \frac{5}{4}$.

Действительная ось $2a=8$; мнимая ось $2b=6$.

Уравнения директрис: $x = \pm \frac{16}{5}$.

Уравнения асимптот: $y = \pm \frac{3}{4}x$.

Задача 2. Составить уравнение эллипса, зная, что он проходит через точки $M_1(2, 3)$ и $M_2\left(1, \frac{3\sqrt{5}}{2}\right)$.

Решение. Возьмем каноническое уравнение эллипса $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ и вместо текущих координат подставим сначала координаты точки M_1 , а затем координаты точки M_2 . Из полученной системы уравнений:

$$\frac{4}{a^2} + \frac{9}{b^2} = 1; \quad \frac{1}{a^2} + \frac{45}{4b^2} = 1$$

определим параметры эллипса a и b . Обозначив

$$\frac{1}{a^2} = m, \quad \frac{1}{b^2} = n,$$

сведем данную систему к следующей системе:

$$4m + 9n = 1; \quad m + \frac{45}{4n} = 1.$$

Решая ее, получим, что

$$m = \frac{1}{16}, \quad n = \frac{1}{12},$$

откуда $a^2 = 16$, $b^2 = 12$.

Следовательно, искомое уравнение эллипса будет $\frac{x^2}{16} + \frac{y^2}{12} = 1$.

Задача 3. Дана гипербола $x^2 - y^2 = 8$. Найти софокусный эллипс (т. е. эллипс, фокусы которого совпадают с фокусами данной гиперболы), проходящий через точку $M(4, 6)$.

Решение. Обозначим параметры данной гиперболы через a_1 , b_1 , c_1 и найдем их, приведя уравнение гиперболы к каноническому виду:

$$\frac{x^2}{8} - \frac{y^2}{8} = 1; \quad a_1^2 = 8, \quad b_1^2 = 8, \quad c_1 = \sqrt{8+8} = 4.$$

Так как по условию фокусы искомого эллипса и данной гиперболы совпадают, то для эллипса $c = c_1 = 4$.

Рассмотрим каноническое уравнение эллипса $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ и подставим в него координаты точки $M(4, 6)$: $\frac{16}{a^2} + \frac{36}{b^2} = 1$. У эллипса $a^2 = b^2 + c^2$ или $a^2 = b^2 + 16$.

Подставляя это значение в уравнение эллипса, получим: $\frac{16}{b^2 + 16} + \frac{36}{b^2} = 1$,
или

$$b^2 = 18 \pm 30; \quad b^2 = 48; \quad (b^2 \neq -12), \\ a^2 = b^2 + c^2 = 64.$$

Искомое уравнение эллипса будет иметь вид: $\frac{x^2}{64} + \frac{y^2}{48} = 1$.

Задача 4. Уравнение эллипса $\frac{x^2}{9} + \frac{y^2}{5} = 1$ привести к уравнению в полярной системе координат вида $\rho = \frac{a}{1 - \varepsilon \cos \varphi}$.

Решение. Найдем из данного уравнения параметры a , b , c , затем найдем эксцентриситет $\varepsilon = \frac{c}{a}$ и фокальный параметр эллипса $p = \frac{b^2}{a}$, $a^2 = 9$, $b^2 = 5$, $c^2 = 4$, $\varepsilon = \frac{2}{3}$, $p = \frac{5}{3}$.

Искомое уравнение будет иметь вид

$$\rho = \frac{\frac{5}{3}}{1 - \frac{2}{3} \cos \varphi}$$

или $\rho = \frac{5}{3 - 2 \cos \varphi}$.

Задача 5. Дано уравнение кривой в полярных координатах $\rho = \frac{3}{1 - 2 \cos \varphi}$. Привести его к каноническому уравнению в прямоугольных координатах.

Решение. В данном уравнении $e = \frac{c}{a} = 2$, $p = \frac{b^2}{a} = 3$.

Так как эксцентриситет $e > 1$, то данное уравнение является уравнением гиперболы, у которой $b^2 = c^2 - a^2$. Таким образом, данные параметры могут быть записаны в виде системы двух уравнений $\frac{c}{a} = 2$, $\frac{c^2 - a^2}{a} = 3$. Из этой системы находим, что $a = 1$, $b^2 = 3$. Следовательно, искомое уравнение гиперболы имеет вид:

$$\frac{x^2}{1} - \frac{y^2}{3} = 1.$$

Раздел 5

ОБЩЕЕ УРАВНЕНИЕ ЛИНИЙ ВТОРОГО ПОРЯДКА

§ 29. Уравнения кривых второго порядка с осями симметрии, параллельными координатным осям

Поставим задачу: составить уравнение эллипса с центром в точке $O(m, n)$, оси симметрии которого параллельны координатным осям OX и OY и полуоси соответственно равны a и b (рис. 65).

Возьмем новую систему координат с началом координат в точке $O'(m, n)$ с осями $O'X'$, $O'Y'$, параллельными соответственно осям OX , OY . Тогда уравнение данного эллипса относительно новой системы координат может быть записано в канонической форме:

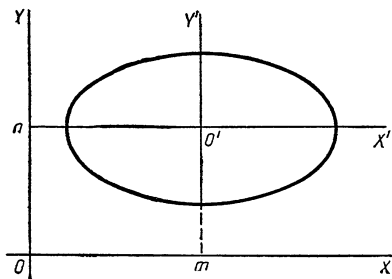


Рис. 65.

$$\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1. \quad (1)$$

Так как $x' = x - m$ и $y' = y - n$, то в старой системе координат уравнение эллипса запишется в виде

$$\frac{(x - m)^2}{a^2} + \frac{(y - n)^2}{b^2} = 1. \quad (2)$$

Уравнение (2) является уравнением эллипса с полуосями a и b , центром в точке $O'(m, n)$ и осями симметрии, параллельными координатным осям.

Решая аналогичным образом задачу относительно уравнения гиперболы, центр которой находится в точке $O'(m, n)$, оси симметрии параллельны координатным осям, а полуоси равны a (действительная), b (мнимая), получим следующее уравнение:

$$\frac{(x - m)^2}{a^2} - \frac{(y - n)^2}{b^2} = 1. \quad (3)$$

Аналогично найдем, что уравнение параболы, ось симметрии которой параллельна оси абсцисс, вершина лежит в точке $O'(m, n)$ и фокальный параметр которой равен p , имеет вид:

$$(y - n)^2 = 2p(x - m). \quad (4)$$

Если ось параболы параллельна оси ординат, то аналогичное преобразование уравнения параболы приведет к уравнению

$$(x - m)^2 = 2p(y - n). \quad (5)$$

§ 30. Исследование общего уравнения линии второго порядка, не содержащего члена с произведением текущих координат

Общее уравнение второй степени с неизвестными x и y имеет вид:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \quad (1)$$

Если $B = 0$ ¹⁾, то уравнение (1) записывается в виде:

$$Ax^2 + Cy^2 + 2Dx + 2Ey + F = 0. \quad (2)$$

В процессе исследования уравнения (2) рассмотрим следующие случаи:

1) $A > 0, C > 0$ (случай $A < 0, C < 0$ сводится к $A > 0, C > 0$ путем изменения знаков перед всеми членами уравнения).

Преобразуем уравнение, дополнив до полного квадрата члены, содержащие x^2 и x , и члены, содержащие y^2 и y :

$$\begin{aligned} Ax^2 + 2Dx + \frac{D^2}{A} + Cy^2 + 2Ey + \frac{E^2}{C} + F &= \frac{D^2}{A} + \frac{E^2}{C}, \\ A \left(x^2 + 2 \frac{D}{A} x + \frac{D^2}{A^2} \right) + C \left(y^2 + 2 \frac{E}{C} y + \frac{E^2}{C^2} \right) &= \frac{D^2}{A} + \frac{E^2}{C} - F, \\ A \left(x + \frac{D}{A} \right)^2 + C \left(y + \frac{E}{C} \right)^2 &= \frac{D^2}{A} + \frac{E^2}{C} - F. \end{aligned}$$

Рассмотрим сначала случай, когда $\frac{D^2}{A} + \frac{E^2}{C} - F = 0$. Уравнение $A \left(x + \frac{D}{A} \right)^2 + C \left(y + \frac{E}{C} \right)^2 = 0$ является уравнением кривой второго порядка, вырождающейся в точку

$$O' \left(-\frac{D}{A}, -\frac{E}{C} \right).$$

Если

$$\frac{D^2}{A} + \frac{E^2}{C} - F < 0,$$

то исследуемому уравнению не удовлетворяет ни одной точки.

Рассмотрим, наконец, случай, когда $\frac{D^2}{A} + \frac{E^2}{C} - F > 0$. Тогда исследуемое уравнение можно переписать так:

$$\frac{A \left(x + \frac{D}{A} \right)^2}{\frac{D^2}{A} + \frac{E^2}{C} - F} + \frac{C \left(y + \frac{E}{C} \right)^2}{\frac{D^2}{A} + \frac{E^2}{C} - F} = 1$$

или

$$\frac{\left(x + \frac{D}{A} \right)^2}{\frac{D^2}{A^2} + \frac{E^2}{AC} - \frac{F}{A}} + \frac{\left(y + \frac{E}{C} \right)^2}{\frac{D^2}{AC} + \frac{E^2}{C^2} - \frac{F}{C}} = 1. \quad (3)$$

Сопоставив уравнение (3) с уравнением эллипса [(2), § 29], заключаем, что в случае, когда $\frac{D^2}{A} + \frac{E^2}{C} - F > 0$, исследуемое уравнение

¹⁾ См. стр. 35.

²⁾ Случай $B \neq 0$ будет рассмотрен в § 31.

является уравнением эллипса. Центр этого эллипса находится в точке $O'(m, n)$, $m = -\frac{D}{A}$, $n = -\frac{E}{C}$, оси симметрии параллельны координатным осям и квадраты полуосей соответственно равны:

$$a^2 = \frac{D^2}{A^2} + \frac{E^2}{AC} - \frac{F}{A}, \quad b^2 = \frac{D^2}{AC} + \frac{E^2}{C^2} - \frac{F}{C}.$$

Таким образом, кривая второго порядка, определяемая уравнением $Ax^2 + Cy^2 + 2Dx + 2Ey + F = 0$, в котором коэффициенты A и C одного знака, либо является эллипсом (действительным или мнимым), либо вырождается в точку.

2) $A > 0$, $C < 0$ (случай $A < 0$, $C > 0$ сводится к $A > 0$, $C < 0$ путем изменения знаков перед всеми членами уравнения).

Повторяя преобразования, сделанные выше, приведем уравнение (2) к виду:

$$A \left(x + \frac{D}{A} \right)^2 - |C| \left(y + \frac{E}{C} \right)^2 = \frac{D^2}{A} + \frac{E^2}{C} - F.$$

Если $\frac{D^2}{A} + \frac{E^2}{C} - F = 0$, то уравнение

$$A \left(x + \frac{D}{A} \right)^2 - |C| \left(y + \frac{E}{C} \right)^2 = 0$$

есть уравнение кривой второго порядка, распадающейся на две пересекающиеся прямые.

Если же $\frac{D^2}{A} + \frac{E^2}{C} - F \neq 0$, то уравнение (2) приводится к виду:

$$\frac{\left(x + \frac{D}{A} \right)^2}{\frac{D^2}{A^2} + \frac{E^2}{AC} - \frac{F}{A}} - \frac{\left(y + \frac{E}{C} \right)^2}{\frac{D^2}{A|C|} + \frac{E^2}{C|C|} - \frac{F}{|C|}} = 1. \quad (4)$$

Если $\frac{D^2}{A} + \frac{E^2}{C} - F > 0$, то уравнение (4) является уравнением гиперболы, центр которой находится в точке $O'(m, n)$, $m = -\frac{D}{A}$, $n = -\frac{E}{C}$, оси симметрии параллельны координатным осям, а квадраты полуосей соответственно равны:

$$a^2 = \frac{D^2}{A^2} + \frac{E^2}{AC} - \frac{F}{A}, \quad b^2 = \frac{D^2}{A|C|} + \frac{E^2}{C|C|} - \frac{F}{|C|} \quad (\text{см. § 29, 3}).$$

Аналогично показывается, что в случае $\frac{D^2}{A} + \frac{E^2}{C} - F < 0$ уравнение (4) также будет уравнением некоторой гиперболы.

Таким образом, кривая второго порядка, определяемая уравнением $Ax^2 + Cy^2 + 2Dx + 2Ey + F = 0$, в котором коэффициенты A и C разного знака, либо является гиперболой, либо распадается на пару пересекающихся прямых.

3) $A > 0$, $C = 0$ (случай $A < 0$, $C = 0$ сводится к этому). Уравнение (2) имеет вид:

$$Ax^2 + 2Dx + 2Ey + F = 0. \quad (5)$$

Преобразуя его, получим $Ax^2 + 2Dx + \frac{D^2}{A} = -2Ey - F + \frac{D^2}{A}$ или, если $E \neq 0$,

$$\begin{aligned} A \left(x + \frac{D}{A} \right)^2 &= -2E \left(y + \frac{F}{2E} - \frac{D^2}{2AE} \right), \\ \left(x + \frac{D}{A} \right)^2 &= -\frac{2E}{A} \left(y + \frac{F}{2E} - \frac{D^2}{2AE} \right). \end{aligned} \quad (6)$$

Уравнение (6) представляет собой уравнение параболы с вершиной в точке $O'(m, n)$, где $m = -\frac{D}{A}$, $n = -\frac{F}{2E} + \frac{D^2}{2AE}$, с осью симметрии, параллельной оси ординат, и «ветвями», направленными вверх, если $-\frac{2E}{A} > 0$, и вниз, если $-\frac{2E}{A} < 0$.

Если $E = 0$, то кривая второго порядка имеет уравнение $Ax^2 + 2Dx + F = 0$, т. е.

$$\left(x + \frac{D}{A}\right)^2 - \left(\frac{D^2}{A^2} - \frac{F}{A}\right) = 0. \quad (7)$$

Если $\frac{D^2}{A^2} - \frac{F}{A} \geq 0$, то кривая распадается на пару параллельных оси OY прямых, имеющих уравнения

$$x = -\frac{D}{A} \pm \sqrt{\frac{D^2}{A^2} - \frac{F}{A}}$$

(при $\frac{D^2}{A^2} - \frac{F}{A} = 0$ эти прямые совпадают).

Если $\frac{D^2}{A^2} - \frac{F}{A} < 0$, то уравнению (7) никакого геометрического места точек на плоскости не соответствует (пара мнимых прямых).

4) $A = 0$, $C > 0$, (случай $A = 0$, $C < 0$ сводится к этому), уравнение (2) примет вид:

$$Cy^2 + 2Dx + 2Ey + F = 0. \quad (8)$$

Преобразуя его, получим (если $D \neq 0$):

$$\left(y + \frac{E}{C}\right)^2 = -2\frac{D}{C}\left(x + \frac{F}{2D} - \frac{E^2}{2DC}\right). \quad (9)$$

Уравнение (9) является уравнением параболы с вершиной в точке $O'(m, n)$, где $m = -\frac{F}{2D} + \frac{E^2}{2DC}$, $n = -\frac{E}{C}$, с осью симметрии, параллельной оси абсцисс, и «ветвями», направленными вправо, если $-2\frac{D}{C} > 0$, и влево, если $-2\frac{D}{C} < 0$. В случае, когда $D = 0$, кривая второго порядка имеет уравнение $Cy^2 + 2Ey + F = 0$, т. е. уравнение

$$\left(y + \frac{E}{C}\right)^2 - \left(\frac{E^2}{C^2} - \frac{F}{C}\right) = 0.$$

Если $\frac{E^2}{C^2} - \frac{F}{C} \geq 0$, то кривая распадается на пару параллельных оси OY прямых¹⁾, имеющих уравнения $y = -\frac{E}{C} \pm \sqrt{\frac{E^2}{C^2} - \frac{F}{C}}$.

Если $\frac{E^2}{C^2} - \frac{F}{C} < 0$, то данная кривая второго порядка является мнимой кривой (будем говорить, что она распадается на пару мнимых прямых).

Таким образом, кривая второго порядка, определяемая уравнением $Ax^2 + Cy^2 + 2Dx + 2Ey + F = 0$, в котором или A или C равно нулю, либо является параболой, либо распадается на пару параллельных (или совпадающих) прямых, либо является парой мнимых прямых²⁾.

Пример. Определить расположение и вид следующих кривых второго порядка:

$$1) 4x^2 + 9y^2 - 16x + 18y - 11 = 0, 4x^2 - 16x + 16 + 9y^2 + 18y + 9 = 11 + 16 + 9, \\ 4(x-2)^2 + 9(y+1)^2 = 36, \frac{(x-2)^2}{9} + \frac{(y+1)^2}{4} = 1.$$

¹⁾ В частности, если $\frac{E^2}{C^2} - \frac{F}{C} = 0$, то эти параллельные прямые совпадают.

²⁾ Случай, когда $A = C = 0$, не рассматриваем, так как при этом уравнение (2) перестает быть уравнением второй степени.

Это уравнение эллипса, центр которого лежит в точке $M(2, -1)$, а полуоси равны: $a=3, b=2$.

$$\begin{aligned} 2) \quad & 2x^2 + 3y^2 + 16x - 18y + 59 = 0, \\ & 2x^2 + 16x + 32 + 3y^2 - 18y + 27 = 0, \\ & 2(x+4)^2 + 3(y-3)^2 = 0, \\ & \frac{(x+4)^2}{5} + \frac{(y-3)^2}{3} = 0. \end{aligned}$$

Это уравнение является уравнением кривой, выродившейся в точку $M(-4, 3)$.

$$\begin{aligned} 3) \quad & 3x^2 - 5y^2 - 18x + 20y - 8 = 0, \\ & 3x^2 - 18x + 27 - 5y^2 + 20y - 20 = 15, \\ & 3(x-3)^2 - 5(y-2)^2 = 15, \\ & \frac{(x-3)^2}{5} - \frac{(y-2)^2}{3} = 1. \end{aligned}$$

Уравнение является уравнением гиперболы с центром в точке $M(3, 2)$ и полуосями, равными: $a = \sqrt{5}, b = \sqrt{3}$.

$$\begin{aligned} 4) \quad & x^2 - y^2 + 10x + 2y + 24 = 0, \\ & x^2 + 10x + 25 - y^2 + 2y - 1 = 0, \\ & (x+5)^2 - (y-1)^2 = 0. \end{aligned}$$

Это уравнение является уравнением пары прямых $y = -x - 4, y = x + 6$, пересекающихся в точке $M(-5, 1)$.

$$\begin{aligned} 5) \quad & y^2 - 6y + 6x + 33 = 0, \\ & y^2 - 6y + 9 = -6x - 24, \\ & (y-3)^2 = -6(x+4). \end{aligned}$$

Это уравнение является уравнением параболы с осью симметрии, параллельной оси абсцисс, с ветвями, направленными влево, с вершиной $M(-4, 3)$ и фокальным параметром $p=3$.

$$\begin{aligned} 6) \quad & x^2 + 10x - 3y + 31 = 0, \\ & x^2 + 10x + 25 = 3y - 6, \\ & (x+5)^2 = 3(y-2). \end{aligned}$$

Полученное уравнение является уравнением параболы с осью симметрии, параллельной оси ординат, с ветвями, направленными вверх, с вершиной $M(-5, 2)$ и фокальным параметром $p = \frac{3}{2}$.

§ 31. Упрощение общего уравнения линий второго порядка

Рассмотрим теперь общее уравнение второй степени с двумя неизвестными:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \quad (1)$$

где $B \neq 0$.

Постараемся так преобразовать это уравнение, чтобы в нем исчез член с произведением переменных. Этого можно добиться надлежащим поворотом координатных осей. Повернем координатные оси на некоторый угол α , величину которого обусловим впоследствии. Используя формулы поворота осей (см. § 8, 4), выразим старые координаты (x, y) через новые (x', y') :

$$\begin{aligned} A(x' \cos \alpha - y' \sin \alpha)^2 + 2B(x' \cos \alpha - y' \sin \alpha)(x' \sin \alpha + y' \cos \alpha) + \\ + C(x' \sin \alpha + y' \cos \alpha)^2 + 2D(x' \cos \alpha - y' \sin \alpha) + \\ + 2E(x' \sin \alpha + y' \cos \alpha) + F = 0. \end{aligned}$$

Раскрыв скобки и сделав приведение подобных членов, обозначим коэффициент при x'^2 через A_1 , коэффициент при $x'y'$ — через $2B_1$ и т. д.

Тогда уравнение (1) в новой системе координат примет вид:

$$A_1x'^2 + 2B_1x'y' + C_1y'^2 + 2D_1x' + 2E_1y' + F = 0. \quad (2)$$

Подберем теперь угол α так, чтобы коэффициент B_1 обратился в нуль, т. е., чтобы $(C - A) \sin \alpha \cdot \cos \alpha + B(\cos^2 \alpha - \sin^2 \alpha) = 0$, т. е. $(C - A) \sin 2\alpha + 2B \cos 2\alpha = 0$, т. е. $(C - A) + 2B \operatorname{ctg} 2\alpha = 0$ (учитывая, что $\sin 2\alpha \neq 0$, так как $B \neq 0$).

Отсюда

$$\operatorname{ctg} 2\alpha = \frac{A-C}{2B}. \quad (3)$$

При повороте координатных осей на угол α , удовлетворяющий (3), уравнение (1) приводится к уравнению вида (2) из § 30.

Все изложенное показывает, что кривые второго порядка (если не считать случаев распада и вырождения) могут быть только эллипсами (в частности, окружностями), гиперболами и парабололами.

Пример. Составить уравнение равнобочной гиперболы, для которой асимптотами являются прямые $x=m$, $y=n$, параллельные координатным осям.

Решение. Уравнение равнобочной гиперболы, для которой координатные оси OX и OY являются асимптотами, имеет вид: $xy=c$. Уравнение же равнобочной гиперболы, для которой асимптотами являются прямые $x=m$, $y=n$, параллельные координатным осям и пересекающиеся, очевидно, в точке $O'(m, n)$, являющейся центром гиперболы, может быть выведено с помощью формул параллельного переноса координатных осей (рис. 66).

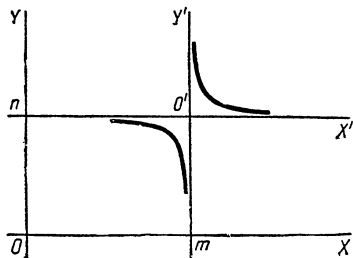


Рис. 66.

Построим новую систему координат, беря за начало точку $O'(m, n)$ и за оси $O'X'$, $O'Y'$ асимптоты гиперболы.

Тогда по отношению к новой системе координат уравнение гиперболы будет иметь вид $x'y'=c$.

При переходе к старой системе координат уравнение (4) преобразуется к виду:

$$(x-m)(y-n)=c \quad (5)$$

или

$$xy - nx - my + mn - c = 0. \quad (6)$$

Раздел 6

ОПРЕДЕЛИТЕЛИ И ИХ ПРИМЕНЕНИЕ К РЕШЕНИЮ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

§ 32. Определители второго порядка и их свойства

Пусть дана квадратная таблица из четырех чисел

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}. \quad (1)$$

Будем называть такую таблицу *матрицей*, а составляющие ее числа a_1 , b_1 , a_2 , b_2 — *элементами* этой матрицы.

Определителем (или, иначе, *детерминантом*) *второго порядка*, соответствующим матрице (1), называют число D , равное $a_1b_2 - a_2b_1$.

Обозначается определитель D символом

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Таким образом,

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1.$$

Например,

$$\begin{vmatrix} -1 & 2 \\ 3 & 4 \end{vmatrix} = (-1) \cdot 4 - 3 \cdot 2 = -10, \quad \begin{vmatrix} 5 & 0 \\ -1 & 3 \end{vmatrix} = 5 \cdot 3 - (-1) \cdot 0 = 15 \text{ и т. д.}$$

Элементы a_1 , b_1 , a_2 и b_2 матрицы (1) называются также *элементами определителя* D . При этом принято говорить, что элементы a_1 и b_1 составляют первую строку, элементы a_2 и b_2 — вторую строку определителя D , а элементы a_1 , a_2 и b_1 , b_2 — соответственно первый и второй столбцы этого определителя. Принято также говорить, что элементы a_1 и b_2 лежат на левой (идущей слева вниз), а элементы b_1 и a_2 на правой (идущей справа вниз) диагонали определителя D .

Пользуясь этой терминологией, можно сказать, что определитель второго порядка равен разности произведений элементов, находящихся на левой и правой диагоналях.

Рассмотрим простейшие свойства определителей второго порядка.

Свойство 1. Определитель не изменится, если его строки заменить столбцами и наоборот.

Действительно, пусть

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad D_1 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Так как

$$D = a_1 b_2 - a_2 b_1,$$

а

$$D_1 = a_1 b_2 - b_1 a_2 = a_1 b_2 - a_2 b_1,$$

то

$$D = D_1.$$

Примечание. Так как строки определителя можно, не меняя его значения, заменить столбцами, то, доказав справедливость некоторого предложения для столбцов определителя, мы тем самым докажем его и для строк. Исходя из этих соображений рассматриваемые ниже свойства определителей будут устанавливаться только для столбцов определителя (или только для его строк).

Свойство 2. При перестановке двух столбцов или двух строк определитель меняет только знак.

Доказательство. Пусть

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad D_1 = \begin{vmatrix} b_1 & a_1 \\ b_2 & a_2 \end{vmatrix}.$$

Тогда

$$D = a_1 b_2 - a_2 b_1, \quad D_1 = b_1 a_2 - b_2 a_1 = -(a_1 b_2 - a_2 b_1)$$

т. е.

$$D_1 = -D.$$

Свойство 3. Определитель, имеющий два одинаковых столбца (строки), равен нулю.

Действительно, если

$$D = \begin{vmatrix} a_1 & a_1 \\ a_2 & a_2 \end{vmatrix},$$

то

$$D = a_1 a_2 - a_1 a_2 = 0.$$

Свойство 4. Если все элементы какого-либо столбца (строки) определителя умножить на одно и то же число m , то определитель умножится на это число.

Доказательство. Пусть, например,

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad D_1 = \begin{vmatrix} ma_1 & b_1 \\ ma_2 & b_2 \end{vmatrix}.$$

Тогда

$$D_1 = (ma_1) b_2 - (ma_2) b_1 = m(a_1 b_2 - a_2 b_1) = mD.$$

Свойство 4 иногда формулируют следующим образом. Если все элементы какого-либо столбца (строки) определителя содержат общий множитель m , то его можно вынести за знак определителя. Так, например, по доказанному

$$\begin{vmatrix} ma_1 & b_1 \\ ma_2 & b_2 \end{vmatrix} = m \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Свойство 5. Определитель, у которого элементы двух столбцов (строк) соответственно пропорциональны, равен нулю.

Действительно, пусть, например,

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix},$$

причем $a_1 = mb_1$, $a_2 = mb_2$.

Тогда, на основании свойств 4 и 3, имеем:

$$D = \begin{vmatrix} mb_1 & b_1 \\ mb_2 & b_2 \end{vmatrix} = m \begin{vmatrix} b_1 & b_1 \\ b_2 & b_2 \end{vmatrix} = 0.$$

Свойство 6. Если каждый элемент какого-либо столбца (строки) определителя есть сумма двух слагаемых, то определитель равен сумме двух определителей, у одного из них элементами соответствующего столбца (строки) являются первые слагаемые, у другого — вторые. Остальные элементы у этих двух определителей те же, что и у данного.

Пусть, например,

$$D = \begin{vmatrix} a'_1 + a''_1 & b_1 \\ a'_2 + a''_2 & b_2 \end{vmatrix}, \quad D_1 = \begin{vmatrix} a'_1 & b_1 \\ a'_2 & b_2 \end{vmatrix}, \quad D_2 = \begin{vmatrix} a''_1 & b_1 \\ a''_2 & b_2 \end{vmatrix}.$$

Покажем, что $D = D_1 + D_2$.

Действительно,

$$D = (a'_1 + a''_1)b_2 - (a'_2 + a''_2)b_1 = (a'_1b_2 - a'_2b_1) + (a''_1b_2 - a''_2b_1) = D_1 + D_2.$$

Свойства 4 и 6 выражают соответственно правило *умножения определителя на число* и правило *сложения определителей*.

Свойство 7. Определитель не изменится, если к элементам какого-либо его столбца (строки) прибавить соответствующие элементы другого столбца (строки), умноженные на одно и то же число.

Доказательство. Пусть, например,

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad D_1 = \begin{vmatrix} a_1 + mb_1 & b_1 \\ a_2 + mb_2 & b_2 \end{vmatrix}.$$

Покажем, что $D = D_1$.

Действительно, на основании свойств 6 и 5, имеем:

$$D_1 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \begin{vmatrix} mb_1 & b_1 \\ mb_2 & b_2 \end{vmatrix} = D + 0 = D.$$

§ 33. Определители третьего порядка и их свойства

Рассмотрим квадратную таблицу (*матрицу*) из девяти чисел (*элементов матрицы*):

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}. \quad (1)$$

Определителем третьего порядка, соответствующим матрице (1), называется число D , равное следующей алгебраической сумме произведений элементов матрицы:

$$a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 - a_1c_2b_3 - b_1a_2c_3 \quad (2)$$

Обозначается этот определитель третьего порядка символом

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Таким образом, по определению

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - c_1 b_2 a_3 - a_1 c_2 b_3 - b_1 a_2 c_3.$$

Числа $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3$ называются *элементами определителя*. Элементы a_1, b_1, c_1 составляют первую строку определителя D , элементы a_2, b_2, c_2 и a_3, b_3, c_3 — вторую и третью его строки, элементы $a_1, a_2, a_3; b_1, b_2, b_3$ и c_1, c_2, c_3 составляют соответственно первый, второй и третий столбцы определителя D . Будем говорить, что элементы a_1, b_2, c_3 лежат на главной левой, а элементы c_1, b_2, a_3 — на главной правой диагонали определителя.

Несмотря на то, что выражение определителя третьего порядка через элементы матрицы (1) является несколько громоздким, закон его составления оказывается весьма простым. Действительно, в выражение (2) со знаком плюс входят произведения элементов, лежащих по главной левой диагонали, а также произведения элементов, лежащих на параллелях к этой диагонали, с добавлением третьего множителя из противоположного угла матрицы (см. схему 1), а со знаком минус — произведение

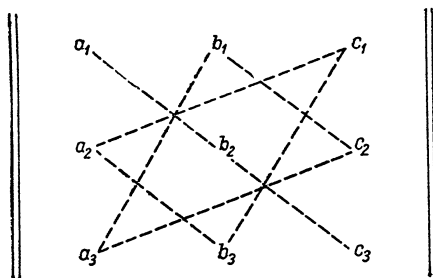


Схема 1

элементов, лежащих на главной правой диагонали, а также произведения элементов, лежащих на параллелях правой диагонали с добавлением третьего множителя из противоположного угла матрицы (см. схему 2).

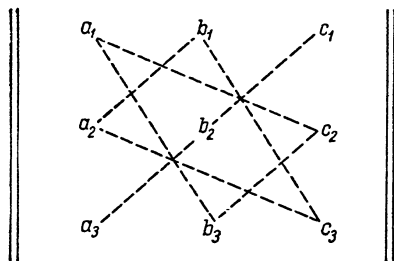


Схема 2

Это правило, называемое *правилом треугольников* или иначе, *правилом Саррюса*, позволяет сравнительно просто вычислять определители третьего порядка.

Пример 1.

$$\begin{vmatrix} 1 & -1 & 2 \\ 3 & -1 & 1 \\ 2 & 1 & 4 \end{vmatrix} = 1 \cdot (-1) \cdot 4 + (-1) \cdot 1 \cdot 2 + 3 \cdot 1 \cdot 2 - 2 \cdot (-1) \cdot 2 - 1 \cdot 1 \cdot 1 - \\ - 3 \cdot (-1) \cdot 4 = -4 - 2 + 6 + 4 - 1 + 12 = 15.$$

Пример 2.

$$\begin{vmatrix} 0 & 2 & 5 \\ -1 & 0 & 7 \\ -2 & 0 & 8 \end{vmatrix} = 0 \cdot 0 \cdot 8 + 2 \cdot 7 \cdot (-2) + 5 \cdot (-1) \cdot 0 - 5 \cdot (-2) \cdot 0 - \\ - 2 \cdot (-1) \cdot 8 - 0 \cdot 7 \cdot 0 = -28 + 16 = -12.$$

Рассмотрим снова матрицу

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

и соответствующий ей определитель

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Минором любого элемента этого определителя называется определитель второго порядка, соответствующий матрице, полученной из данной вычеркиванием строки и столбца, в которых лежит указанный элемент.

Так, например, минором элемента a_2 определителя D , лежащего во второй строке и первом столбце, является определитель

$$\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix},$$

соответствующий матрице, полученной из данной вычеркиванием второй строки и первого столбца.

Для краткости обычно говорят, что минор данного элемента определителя D есть определитель, полученный из определителя D вычеркиванием строки и столбца, в которых лежит этот элемент.

Алгебраическим дополнением любого элемента определителя D называют минор этого элемента, взятый со знаком $(-1)^{i+j}$, где i — номер строки, а j — номер столбца определителя D , в которых лежит данный элемент.

Обозначают алгебраическое дополнение элемента обычно большой буквой того же наименования и с тем же номером (индексом), что и у данного элемента.

Например, алгебраическое дополнение элемента a_1 обозначают A_1 , элемента b_3 — B_3 и т. д.

Таким образом, по определению

$$A_1 = (-1)^{1+1} \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}, \quad B_3 = (-1)^{3+2} \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

и т. д.

Теорема. Определитель D равен сумме произведений элементов любого его столбца или строки на их алгебраические дополнения.

Другими словами, имеют место следующие равенства:

$$\begin{aligned} D &= a_1 A_1 + a_2 A_2 + a_3 A_3, & D &= a_1 A_1 + b_1 B_1 + c_1 C_1, \\ D &= b_1 B_1 + b_2 B_2 + b_3 B_3, & D &= a_2 A_2 + b_2 B_2 + c_2 C_2, \\ D &= c_1 C_1 + c_2 C_2 + c_3 C_3, & D &= a_3 A_3 + b_3 B_3 + c_3 C_3. \end{aligned} \quad (3)$$

Докажем, например, первое из равенств (3). Для этого преобразуем правую часть этого равенства:

$$\begin{aligned} a_1 A_1 + a_2 A_2 + a_3 A_3 &= a_1 \cdot (-1)^{1+1} \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} + a_2 \cdot (-1)^{2+1} \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + \\ &+ a_3 \cdot (-1)^{3+1} \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} = a_1 (b_2 c_3 - b_3 c_2) - a_2 (b_1 c_3 - b_3 c_1) + \\ &+ a_3 (b_1 c_2 - b_2 c_1) = a_1 b_2 c_3 - a_1 c_2 b_3 - b_1 a_2 c_3 + c_1 a_2 b_3 + b_1 c_2 a_3 - c_1 b_2 a_3. \end{aligned}$$

Сравнивая полученное выражение с выражением (2) определителя D , заключаем, что $a_1 A_1 + a_2 A_2 + a_3 A_3 = D$.

Аналогично устанавливаются остальные равенства (3).

Запись определителя D по любой из указанных формул (3) называется *разложением этого определителя по элементам* соответствующего столбца или строки. Эти разложения удобно использовать для вычисления определителей третьего порядка.

Например, разложив определитель

$$D = \begin{vmatrix} 7 & 0 & 1 \\ 3 & 5 & 8 \\ -4 & 2 & 3 \end{vmatrix}$$

по элементам первой строки, получим:

$$D = 7 \cdot \begin{vmatrix} 5 & 8 \\ 2 & 3 \end{vmatrix} + 1 \cdot \begin{vmatrix} 3 & 5 \\ -4 & 2 \end{vmatrix} = 7(15 - 16) + (6 + 20) = 19.$$

Для определителей третьего порядка остаются справедливыми свойства 1—7, рассмотренные в § 32 для определителей второго порядка. Доказательства этих свойств непосредственно следуют из теоремы о разложении определителя третьего порядка по элементам строк и столбцов и соответствующих свойств определителей второго порядка.

Так, например, свойство 1, утверждающее, что определитель не изменится, если его строки заменить столбцами, можно доказать следующим образом.

Пусть

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \quad D_1 = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Раскладывая определитель D по элементам первого столбца:

$$D = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix},$$

а определитель D_1 по элементам первой строки:

$$D_1 = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

и учитывая, что определитель второго порядка от замены строк столбцами не меняется, получим

$$D_1 = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} = D.$$

Аналогично можно установить справедливость остальных свойств. Предоставим это читателю. Свойства определителей, в особенности свойство 7, в сочетании с теоремой о разложении определителя по элементам строк и столбцов, иногда позволяют значительно упростить вычисление определителя.

Пример. Вычислить определитель

$$D = \begin{vmatrix} 3 & 4 & 15 \\ 2 & 8 & 16 \\ 4 & 20 & 7 \end{vmatrix}.$$

Вынося общий множитель элементов второй строки, а затем второго столбца, получаем:

$$D = 8 \begin{vmatrix} 3 & 1 & 15 \\ 1 & 1 & 8 \\ 4 & 5 & 7 \end{vmatrix}.$$

Прибавим теперь к элементам первой строки соответствующие элементы второй строки, умноженные на -1 , а к элементам последней строки соответствующие элементы второй строки, умноженные на -5 ¹⁾. По свойству 7 определитель при этом не изменится.

$$D = 8 \begin{vmatrix} 2 & 0 & 7 \\ 1 & 1 & 8 \\ -1 & 0 & -33 \end{vmatrix}.$$

Разложив полученный определитель по элементам второго столбца, будем иметь:

$$D = 8 \cdot 1 \cdot (-1)^{2+2} \begin{vmatrix} 2 & 7 \\ -1 & -33 \end{vmatrix} = 8(-66 + 7) = -472.$$

§ 34. Применение определителей к исследованию и решению систем линейных уравнений

1. Система двух линейных уравнений с двумя неизвестными. Рассмотрим систему двух уравнений первой степени (*линейных уравнений*) с неизвестными x и y :

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2. \end{cases} \quad (1)$$

Числа a_1 , b_1 , a_2 и b_2 называются *коэффициентами*, а числа c_1 и c_2 — *свободными членами*, или *правыми частями* системы (1). В случае, если $c_1 = c_2 = 0$, система (1) называется *однородной системой линейных уравнений*, в противном случае, т. е. если хотя бы одно из чисел c_1 и c_2 отлично от нуля, — *неоднородной*.

Решением системы (1) является всякая пара чисел x , y , обращающая оба уравнения системы в тождества.

Покажем, как применяются определители второго порядка для исследования и нахождения решений систем (1).

Умножим обе части первого уравнения системы (1) на b_2 , а второго — на b_1 и сложим полученные уравнения. При этом неизвестное y исключается и остается уравнение

$$(a_1b_2 - a_2b_1)x = b_2c_1 - b_1c_2. \quad (2)$$

Аналогично, исключая неизвестное x , приходим к уравнению

$$(a_1b_2 - a_2b_1)y = a_1c_2 - a_2c_1. \quad (3)$$

Разность $a_1b_2 - a_2b_1$ равна определителю $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$, составленному из коэффициентов системы (1). Будем называть его *определителем системы (1)* и обозначать через D . Разности $b_2c_1 - b_1c_2$ и $a_1c_2 - a_2c_1$ равны соответственно определителям

$$\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix},$$

¹⁾ Можно условиться короче говорить так: «вычтем из первой строки вторую, а из третьей строки — пять вторых строк».

полученным из определителя системы (1) заменой столбца коэффициентов в первом случае при x , во втором — при y столбцом свободных членов системы (1).

Обозначим эти определители соответственно через D_x и D_y . В новых обозначениях уравнения (2) и (3) запишутся в виде:

$$\begin{cases} D \cdot x = D_x, \\ D \cdot y = D_y. \end{cases} \quad (4)$$

Если система (1) имеет решение, то это решение удовлетворяет и системе (4), так как система уравнений (4) является следствием системы (1).

Предположим теперь, что определитель D системы (1) отличен от нуля. Тогда, решая систему (4), находим ее единственное решение:

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D}. \quad (5)$$

Подставляя эти значения x и y в систему (1), после несложных выкладок убеждаемся, что определяемые формулами (5) x и y являются решением системы (1).

Таким образом, если определитель системы (1) отличен от нуля, то система имеет, и притом единственное, решение. Его можно найти по формулам

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D},$$

называемым *формулами Крамера*.

Пусть теперь определитель D системы (1) равен нулю. При этом либо оба определителя D_x и D_y также равны нулю, либо хотя бы один из них отличен от нуля.

Если $D = D_x = D_y = 0$, т. е. если $a_1b_1 - a_2b_2 = 0$, $b_2c_1 - b_1c_2 = 0$, $a_1c_2 - a_2c_1 = 0$, то коэффициенты и свободные члены уравнений системы пропорциональны. Это значит, что одно уравнение системы (1), пусть, например, второе, получается из другого умножением на постоянный множитель, т. е. является его следствием. Тогда всякое решение первого уравнения будет решением второго, а значит, и системы (1).

Одно уравнение с двумя неизвестными

$$a_1x + b_1y = c_1$$

имеет бесчисленное множество решений, так как одному неизвестному можно давать произвольные значения, а другое определять так, чтобы удовлетворялось это уравнение. Таким образом, в рассматриваемом случае система (1) имеет бесчисленное множество решений.

Если же $D = 0$, но среди определителей D_x и D_y хотя бы один, пусть D_x , не равен нулю, то система (1) решений не имеет.

Действительно, если $D = 0$, $D_x \neq 0$, то равенство $D \cdot x = D_x$ не может выполняться ни при каких значениях x , в то время как всякое решение системы (1), если бы оно существовало, должно было бы удовлетворять системе (4).

Итак, если определитель системы (1) не равен нулю ($D \neq 0$), то система имеет, и притом единственное, решение; если же определитель системы D равен нулю, то система либо имеет бесчисленное множество решений, либо совсем не имеет решения.

Пример 1. Система $\begin{cases} 2x - 10y = 3, \\ 3x + 3y = 1 \end{cases}$ имеет, и притом единственное, решение, так как $D = \begin{vmatrix} 2 & -10 \\ 3 & 3 \end{vmatrix} = 36 \neq 0$. Находя $D_x = \begin{vmatrix} 3 & -10 \\ 1 & 3 \end{vmatrix} = 19$, $D_y = \begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} = -7$, по формулам Крамера получаем:

$$x = \frac{19}{36}, \quad y = -\frac{7}{36}.$$

Пример 2. Система $\begin{cases} x - 2y = 1, \\ -2x + 4y = -2 \end{cases}$ имеет бесчисленное множество решений, так как $D = \begin{vmatrix} 1 & -2 \\ -2 & 4 \end{vmatrix} = 0$, $D_x = \begin{vmatrix} 1 & -2 \\ -2 & 4 \end{vmatrix} = 0$, $D_y = \begin{vmatrix} 1 & 1 \\ -2 & -2 \end{vmatrix} = 0$. Решения этой системы — пары чисел $x = 1 + 2t$, $y = t$, где t — любое число.

Пример 3. Система $\begin{cases} x + 2y = 1 \\ 3x + 6y = 5 \end{cases}$ решений не имеет, так как $D = \begin{vmatrix} 1 & 2 \\ 3 & 6 \end{vmatrix} = 0$, $D_x = \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} \neq 0$.

Частным случаем системы (1) является *однородная система линейных уравнений*:

$$\begin{cases} a_1x + b_1y = 0, \\ a_2x + b_2y = 0. \end{cases}$$

Очевидно, что всякая однородная система имеет так называемое *нулевое решение*: $x = 0$, $y = 0$.

Учитывая проведенное выше исследование, заключаем, что либо эта система имеет единственное нулевое решение ($D \neq 0$), либо она имеет также бесчисленное множество ненулевых решений ($D = 0$).

2. Однородная система двух линейных уравнений с тремя неизвестными. Однородной системой двух линейных уравнений с неизвестными x , y и z называется система уравнений вида:

$$\begin{cases} a_1x + b_1y + c_1z = 0, \\ a_2x + b_2y + c_2z = 0. \end{cases} \quad (6)$$

Числа a_1 , b_1 , c_1 , a_2 , b_2 , c_2 называются *коэффициентами* такой системы.

Решением системы (6) называется всякая тройка чисел x , y , z , обращающая оба уравнения системы в тождества. Предположим сначала, что хотя бы один из определителей, составленных из коэффициентов при неизвестных:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix},$$

например, определитель $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$, отличен от нуля.

Запишем тогда систему (6) в виде:

$$\begin{cases} a_1x + b_1y = -c_1z, \\ a_2x + b_2y = -c_2z \end{cases} \quad (7)$$

и дадим z произвольное численное значение. Этому значению z будет соответствовать единственное решение системы (7), определяемое, например, по формулам Крамера:

$$\begin{aligned} x &= \frac{\begin{vmatrix} -c_1z & b_1 \\ -c_2z & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z = \frac{\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z, \\ y &= \frac{\begin{vmatrix} a_1 & -c_1z \\ a_2 & -c_2z \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} = -\frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z = \frac{\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z. \end{aligned} \quad (8)$$

Найденные значения x и y вместе с взятым значением z составят решение системы (6). Так как значение z выбиралось произвольно, то

в рассматриваемом случае система (6) имеет бесчисленное множество решений. Совокупность всех решений системы можно записать в виде:

$$\left\{ x = \frac{\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z, \quad y = \frac{\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z, \quad z = z \right\},$$

где z — любое число.

Обозначим отношение $\frac{z}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$ через t . Тогда совокупность решений системы (6) запишется в виде:

$$\left\{ x = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} t, \quad y = \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} t, \quad z = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} t \right\}, \quad (9)$$

где t — множитель пропорциональности — любое число.

В случае, если $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = 0$, но $\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \neq 0$ или $\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} \neq 0$, рассуждения отличаются от проведенных выше лишь переменой ролей неизвестных. Совокупность всех решений системы снова запишется в виде (9).

Пусть теперь все три определителя

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}$$

равны нулю. Тогда коэффициенты уравнений (6) пропорциональны, т. е. одно уравнение системы есть следствие другого. Решение системы (6) сведется к решению одного уравнения с тремя неизвестными. Такое уравнение имеет бесчисленное множество решений. Для их нахождения надо двум неизвестным придавать произвольные значения, а третье выбирать так, чтобы при этом удовлетворялось уравнение.

Пример 1. Система $\begin{cases} 3x - 2y + 7z = 0, \\ x + y - 5z = 0 \end{cases}$ имеет бесчисленное множество решений. Совокупность решений этой системы можно записать в виде:

$$\left\{ x = \begin{vmatrix} -2 & 7 \\ 1 & -5 \end{vmatrix} t, y = \begin{vmatrix} 7 & 3 \\ -5 & 1 \end{vmatrix} t, z = \begin{vmatrix} 3 & -2 \\ 1 & 1 \end{vmatrix} t \right\}$$

или, что то же,

$$\{ x = 3t, y = 22t, z = 5t \},$$

где t — любое число.

Пример 2. Система $\begin{cases} x - 2y - 3z = 0, \\ -2x + 4y + 6z = 0 \end{cases}$ имеет бесчисленное множество решений. Так как второе уравнение системы — есть следствие первого, совокупность всех решений системы можно записать в виде:

$$\{ x = 2t + 3u, y = t, z = u \},$$

где t и u — любые числа.

3. Система трех линейных уравнений с тремя неизвестными. Пусть дана система трех уравнений первой степени с неизвестными x , y и z :

$$\begin{cases} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2, \\ a_3x + b_3y + c_3z = d_3 \end{cases} \quad (10)$$

Числа $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3$ назовем *коэффициентами* системы, а числа d_1, d_2 и d_3 — *свободными членами*, или *правыми частями*, этой системы. В случае, если d_1, d_2 и d_3 равны нулю, система (10) называется *однородной системой линейных уравнений*, в противном случае, т. е. если хотя бы один из свободных членов системы отличен от нуля, — *не-*

однородной. Решением системы (10) является всякая тройка чисел x, y, z , обращающая уравнения системы в тождества. Система, имеющая решение, называется *совместной*.

Поставим своей целью исследовать систему (10), т. е. выяснить, при каких условиях на коэффициенты и свободные члены эта система имеет единственное решение, при каких условиях решений может быть больше одного и когда система вообще не имеет решений.

К исследованию и решению системы (10) можно, так же как и для изученных выше систем уравнений, применять теорию определителей.

Рассмотрим определитель третьего порядка, составленный из коэффициентов при неизвестных x, y и z :

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Будем называть его *определителем системы* (10). Обозначим через $A_1, A_2, A_3, B_1, \dots, C_3$ — соответственно, алгебраические дополнения элементов $a_1, a_2, a_3, b_1, \dots, c_3$ этого определителя.

Умножим обе части первого уравнения системы на A_1 , второго — на A_2 , третьего — на A_3 и сложим полученные уравнения:

$$(a_1 A_1 + a_2 A_2 + a_3 A_3) x + (b_1 A_1 + b_2 A_2 + b_3 A_3) y + (c_1 A_1 + c_2 A_2 + c_3 A_3) z = d_1 A_1 + d_2 A_2 + d_3 A_3. \quad (11)$$

На основании теоремы о разложении определителя по элементам столбцов и строк, заключаем, что

$$a_1 A_1 + a_2 A_2 + a_3 A_3 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = D.$$

Коэффициенты при неизвестных y и z и правая часть уравнения (11) представляют собой разложения по элементам первого столбца определителей, которые получатся из определителя D , если в последнем вместо элементов a_1, a_2 и a_3 взять сначала b_1, b_2, b_3 , затем c_1, c_2 и c_3 и, наконец, d_1, d_2 и d_3 . Так как определители, содержащие одинаковые столбцы, равны нулю, то

$$\begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix} = b_1 A_1 + b_2 A_2 + b_3 A_3 = 0, \\ \begin{vmatrix} c_1 & b_1 & c_1 \\ c_2 & b_2 & c_2 \\ c_3 & b_3 & c_3 \end{vmatrix} = c_1 A_1 + c_2 A_2 + c_3 A_3 = 0.$$

Обозначая $d_1 A_1 + d_2 A_2 + d_3 A_3$ или, что то же, определитель

$$\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}$$

через D_x , запишем уравнение (11) в виде $D \cdot x = D_x$.

Аналогичным образом из второго и третьего уравнений системы (10) получаем:

$$D \cdot y = D_y, \quad D \cdot z = D_z.$$

где

$$D_y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}, \quad D_z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}.$$

Таким образом, можно утверждать, что если система (10) имеет решение, то это решение удовлетворяет также и системе

$$\begin{cases} D \cdot x = D_x, \\ D \cdot y = D_y, \\ D \cdot z = D_z, \end{cases} \quad (12)$$

являющейся следствием данной системы.

Предположим теперь, что определитель данной системы D отличен от нуля ($D \neq 0$). Решая систему (12), находим ее единственное решение:

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D}, \quad z = \frac{D_z}{D}.$$

Подставляя найденные значения x , y и z в уравнения системы (10), после несложных преобразований, использующих свойства определителей, убеждаемся, что каждое из этих уравнений обращается в тождество.

Так, например,

$$\begin{aligned} a_1 \frac{D_x}{D} + b_1 \frac{D_y}{D} + c_1 \frac{D_z}{D} &= \frac{1}{D} [a_1 D_x + b_1 D_y + c_1 D_z] = \\ &= \frac{1}{D} [a_1 (d_1 A_1 + d_2 A_2 + d_3 A_3) + b_1 (d_1 B_1 + d_2 B_2 + d_3 B_3) + \\ &+ c_1 (d_1 C_1 + d_2 C_2 + d_3 C_3)] = \frac{1}{D} [d_1 (a_1 A_1 + b_1 B_1 + c_1 C_1) + \\ &+ d_2 (a_1 A_2 + b_1 B_2 + c_1 C_2) + d_3 (a_1 A_3 + b_1 B_3 + c_1 C_3)]. \end{aligned}$$

Так как

$$\begin{aligned} a_1 A_1 + b_1 B_1 + c_1 C_1 &= D, \\ a_1 A_2 + b_1 B_2 + c_1 C_2 &= 0, \\ a_1 A_3 + b_1 B_3 + c_1 C_3 &= 0^1, \end{aligned}$$

то

$$a_1 \frac{D_x}{D} + b_1 \frac{D_y}{D} + c_1 \frac{D_z}{D} = \frac{1}{D} d_1 D = d_1.$$

Аналогично рассматриваются второе и третье уравнения.

Таким образом, если определитель D системы (10) не равен нулю, то система имеет единственное решение. Это решение можно найти по формулам

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D}, \quad z = \frac{D_z}{D},$$

называемым *формулами Крамера*.

Здесь D — определитель системы, а D_x , D_y и D_z — определители, полученные из D заменой в нем соответственно первого, второго и третьего столбцов столбцом свободных членов данной системы.

Пр и м е р. Решим систему:
$$\begin{cases} x + 2y - z = 1, \\ -3x + y + 2z = 0, \\ x + 4y + 3z = 2. \end{cases}$$

Составляем определитель системы:

$$D = \begin{vmatrix} 1 & 2 & -1 \\ -3 & 1 & 2 \\ 1 & 4 & 3 \end{vmatrix} = 30.$$

¹⁾ Выражения $a_1 A_2 + b_1 B_2 + c_1 C_2$ и $a_1 A_3 + b_1 B_3 + c_1 C_3$ представляют собой разложения по элементам соответственно второй и третьей строк определителей:

$$\begin{vmatrix} a_1 b_1 c_1 \\ a_1 b_1 c_1 \\ a_1 b_1 c_1 \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} a_1 b_1 c_1 \\ a_2 b_2 c_2 \\ a_1 b_1 c_1 \end{vmatrix}.$$

Так как у этих определителей по две строки одинаковые, то они равны нулю.

Так как $D \neq 0$, система имеет единственное решение. Найдем его по формулам Крамера:

$$D_x = \begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 2 & 4 & 3 \end{vmatrix} = 5, \quad D_y = \begin{vmatrix} 1 & 1 & -1 \\ -3 & 0 & 2 \\ 1 & 2 & 3 \end{vmatrix} = 13, \quad D_z = \begin{vmatrix} 1 & 2 & 1 \\ -3 & 1 & 0 \\ 1 & 4 & 2 \end{vmatrix} = 1,$$

следовательно,

$$x = \frac{D_x}{D} = \frac{1}{6}, \quad y = \frac{D_y}{D} = \frac{13}{30}, \quad z = \frac{D_z}{D} = \frac{1}{30}.$$

Пусть теперь определитель D системы равен нулю. Если при этом, хотя бы один из определителей D_x, D_y, D_z отличен от нуля, то данная система решений не имеет.

Действительно, если $D=0$, а, например, $D_x \neq 0$, то равенство $D \cdot x = D_x$ не может выполняться ни при каких значениях x , т. е. система (12) не может иметь решений. Так как всякое решение данной системы должно удовлетворять системе (12), заключаем, что данная система также не может иметь решений.

Пример. Система
$$\begin{cases} x - y + 3z = 1, \\ 2x - 2y + 4z = 3, \\ -x + y - z = 0 \end{cases}$$

несовместна, так как

$$D = \begin{vmatrix} 1 & -1 & 3 \\ 2 & -2 & 4 \\ -1 & 1 & -1 \end{vmatrix} = 0, \quad D_x = \begin{vmatrix} 1 & -1 & 3 \\ 3 & -2 & 4 \\ 0 & 1 & -1 \end{vmatrix} = 4 \neq 0.$$

В том, что эта система не имеет решений, можно убедиться и непосредственно. Действительно, если бы существовало решение системы $\{x_0, y_0, z_0\}$, то, подставляя эти числа в уравнения и вычитая из первого из полученных тождеств сумму второго и третьего, мы получили бы, что $0 = -2$.

Не рассматривая в общем виде систем, у которых $D=0, D_x=0, D_y=0, D_z=0$, заметим, что в этом случае система (10) либо совсем не имеет решений, либо имеет их бесчисленное множество.

Пример 1. Система
$$\begin{cases} x - y + z = 2, \\ 2x - 2y + 2z = 4, \\ 3x - 3y + 3z = 5 \end{cases} \quad (D = D_x = D_y = D_z = 0)$$
 решений не

имеет. Действительно, из первого уравнения следует, что $x - y + z = 2$, а из последнего — что $x - y + z = \frac{5}{3}$, чего не может быть.

Пример 2. Система
$$\begin{cases} x - y + z = 2, \\ 2x - 2y + 2z = 4, \\ 3x - 3y + 3z = 6 \end{cases} \quad (D = D_x = D_y = D_z = 0)$$

имеет бесчисленное множество решений. Совокупность всех ее решений можно записать так: $\{x=t, y=u, z=2-t+u\}$, где t и u — любые числа.

Остановимся подробнее на изучении *однородной системы линейных уравнений*:

$$\begin{cases} a_1x + b_1y + c_1z = 0, \\ a_2x + b_2y + c_2z = 0, \\ a_3x + b_3y + c_3z = 0, \end{cases} \quad (13)$$

являющейся частным случаем системы (10). Очевидно, что всякая однородная система (13) имеет так называемое *нулевое решение*:

$$\{x=0, y=0, z=0\},$$

т. е., другими словами, всякая однородная система линейных уравнений совместна.

В случае, если определитель D системы (13) не равен нулю, нулевое решение будет единственным решением системы. Это решение дают и формулы Крамера, так как у однородной системы $D_x = D_y = D_z = 0$.

Предположим теперь, что определитель D системы (13) равен нулю, но хотя бы один из его миноров отличен от нуля. Пусть, например,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0^1).$$

Рассмотрим первые два уравнения системы (13). По доказанному (см. п. 2), система из этих двух уравнений будет иметь бесчисленное множество решений, причем совокупность всех решений такой системы запишется в виде:

$$\left\{ x = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} t, \quad y = \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} t, \quad z = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} t \right\}, \quad (14)$$

где t — любое число.

Подставляя эти значения x , y и z в третье уравнение системы (13), получаем:

$$\begin{aligned} a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} t + b_3 \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} t + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} t = \\ = t \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = t \cdot D = 0. \end{aligned}$$

Так как всякое решение первых двух уравнений системы (13) в рассматриваемом случае является решением и третьего уравнения, то при решении системы (13) третье уравнение следует отбросить, оно является следствием первых двух.

Пусть теперь $D = 0$ и все миноры определителя D также равны нулю. Тогда соответствующие коэффициенты уравнений системы будут пропорциональны, т. е. каждое из уравнений системы (13) получается из какого-либо одного, пусть, например, первого, умножением на некоторый множитель. Это значит, что решение системы (13) сводится к решению одного линейного уравнения с тремя неизвестными:

$$a_1 x + b_1 y + c_1 z = 0$$

(два других уравнения являются его следствиями).

В рассматриваемом случае система (13) снова имеет бесчисленное множество решений (два неизвестных берутся произвольно, а третье так, чтобы $a_1 x + b_1 y + c_1 z = 0$).

Таким образом, если определитель однородной системы (13) не равен нулю, то система имеет единственное нулевое решение. Если определитель этой системы равен нулю, то система имеет бесчисленное множество решений²⁾. Отсюда, как следствие, получается такое утверждение: для того чтобы однородная система (13) имела ненулевые решения, необходимо и достаточно, чтобы определитель этой системы был равен нулю.

Пример 1. Система $\begin{cases} x - y + z = 0, \\ 2x + 3y - z = 0, \\ x + y + z = 0 \end{cases}$ имеет только нулевое решение $x =$

$= 0, y = 0, z = 0$, так как

$$D = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 3 & -1 \\ 1 & 1 & 1 \end{vmatrix} = 6 \neq 0.$$

¹⁾ Переменив порядок следования уравнений и неизвестных, всегда можно ненулевой минор определителя D получить в левом верхнем углу.

²⁾ Одним из них является нулевое решение системы.

Пример 2. Система $\begin{cases} 2x + y + 3z = 0, \\ 3x + 2y + 5z = 0, \\ 5x + 4y + 9z = 0 \end{cases}$ имеет бесчисленное множество решений, так как

$$D = \begin{vmatrix} 2 & 1 & 3 \\ 3 & 2 & 5 \\ 5 & 4 & 9 \end{vmatrix} = 0.$$

Поскольку минор $\begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix}$ отличен от нуля, третье уравнение является следствием первых двух.

Решая систему $\begin{cases} 2x + y + 3z = 0, \\ 3x + 2y + 5z = 0, \end{cases}$ находим

$$x = -t, y = -t, z = t.$$

Совокупность всех решений данной системы:

$$\{x = -t, y = -t, z = t\},$$

где t — любое число.

Пример 3. Система $\begin{cases} x + y - 2z = 0, \\ 3x + 3y - 6z = 0, \\ 4x + 4y - 8z = 0, \end{cases}$ имеет бесчисленное множество решений, поскольку

$$D = \begin{vmatrix} 1 & 1 & -2 \\ 3 & 3 & -6 \\ 4 & 4 & -8 \end{vmatrix} = 0.$$

Так как все миноры определителя D равны нулю, решение системы сводится к решению одного уравнения:

$$x + y - 2z = 0.$$

Совокупность всех решений данной системы:

$$\left\{ x = t, y = u, z = \frac{t+u}{2} \right\},$$

где t и u — любые числа.

§ 35. Применение определителей к решению некоторых задач аналитической геометрии на плоскости

1. Уравнение прямой, проходящей через две точки. Найдем уравнение прямой, проходящей через две точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$. Для этого запишем искомое уравнение в общем виде: $Ax + By + C = 0$ (коэффициенты A , B и C пока неизвестны).

Пусть $M(x, y)$ — произвольная точка данной прямой. Так как точки M , M_1 и M_2 лежат на этой прямой, то их координаты удовлетворяют уравнению прямой:

$$\begin{aligned} Ax + By + C &= 0, \\ Ax_1 + By_1 + C &= 0, \\ Ax_2 + By_2 + C &= 0. \end{aligned} \tag{1}$$

Будем рассматривать эти соотношения как однородную систему линейных уравнений с неизвестными A , B и C . Для того чтобы система (1) имела ненулевое решение, необходимо и достаточно, чтобы определитель системы был равен нулю:

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0. \tag{2}$$

Уравнение (2), устанавливающее зависимость между координатами произвольной точки $M(x, y)$, линейно относительно x и y (в этом легко убедиться, разложив определитель по элементам первой строки), т. е. является уравнением прямой. Так как, подставляя вместо x, y соответственно x_1, y_1 и x_2, y_2 , мы получаем определители с двумя одинаковыми строками, т. е. равные нулю, то уравнение (2) является уравнением прямой, проходящей через точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$, записанным с помощью определителя.

Пример. Составим уравнение прямой, проходящей через точки $M_1(2, 5)$ и $M_2(3, 7)$. Для этого подставим в уравнение (2) координаты точек M_1 и M_2 :

$$\begin{vmatrix} x & y & 1 \\ 2 & 5 & 1 \\ 3 & 7 & 1 \end{vmatrix} = 0,$$

$$x \begin{vmatrix} 5 & 1 \\ 7 & 1 \end{vmatrix} + y \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix} = 0,$$

т. е.

$$2x - y + 1 = 0.$$

2. Площадь треугольника. Найдем площадь треугольника, зная координаты его вершин: $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ и $M_3(x_3, y_3)$. Искомая площадь $S = \frac{1}{2} M_2 M_3 \cdot h$, где

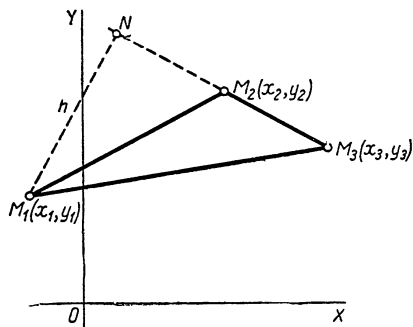


Рис. 67.

$$M_2 M_3 = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}, \quad h = M_1 N, \quad (\text{рис. 67}).$$

Составим уравнение прямой, проходящей через две точки M_2 и M_3 в виде (2)

$$\begin{vmatrix} x & y & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

Разлагая определитель по элементам первой строки, получим:

$$x \begin{vmatrix} y_2 & 1 \\ y_3 & 1 \end{vmatrix} + y \begin{vmatrix} 1 & x_2 \\ 1 & x_3 \end{vmatrix} + \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} = 0$$

или, что то же;

$$x(y_2 - y_3) - y(x_2 - x_3) + (x_2 y_3 - x_3 y_2) = 0.$$

Найдем высоту h как расстояние от точки M_1 до прямой, проходящей через точки M_2 и M_3 :

$$h = \frac{|x_1(y_2 - y_3) - y_1(x_2 - x_3) + (x_2 y_3 - x_3 y_2)|}{\sqrt{(x_2 - x_3)^2 + (y_2 - y_3)^2}}.$$

Тогда

$$S = \frac{1}{2} M_2 M_3 \cdot h = \frac{1}{2} |x_1(y_2 - y_3) - y_1(x_2 - x_3) + (x_2 y_3 - x_3 y_2)|$$

или, что то же,

$$S = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}. \quad (3)$$

3. Условие принадлежности трех точек одной прямой. Для того чтобы три точки $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$ лежали на одной прямой,

необходимо и достаточно, чтобы координаты точки M_1 удовлетворяли уравнению прямой, проведенной через точки M_2 и M_3 . Записывая уравнение этой прямой с помощью определителя в виде:

$$\begin{vmatrix} x & y & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0$$

и подставляя в него вместо x и y соответственно x_1 и y_1 , получим:

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0. \quad (4)$$

Это равенство представляет собой *условие принадлежности трех точек одной прямой*.

4. Условие, при котором три прямые пересекаются в одной точке. Пусть даны уравнения трех прямых:

$$A_1x + B_1y + C_1 = 0, \quad A_2x + B_2y + C_2 = 0, \quad A_3x + B_3y + C_3 = 0.$$

Если точка пересечения этих прямых существует, то ее координаты x_0 , y_0 должны удовлетворять системе трех уравнений:

$$\begin{cases} A_1x + B_1y + C_1 = 0, \\ A_2x + B_2y + C_2 = 0, \\ A_3x + B_3y + C_3 = 0. \end{cases} \quad (5)$$

Будем рассматривать систему (5) как однородную систему линейных уравнений относительно x , y , 1 . Условием существования ненулевых ее решений является равенство нулю определителя системы (5).

Следовательно, условием пересечения трех прямых в одной точке является равенство:

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = 0. \quad (6)$$

§ 36. Определители высшего порядка и их свойства

Определители высшего порядка вводятся как обобщение понятия определителей второго и третьего порядков. Чтобы такое обобщение стало возможным, изучим структуру определителей второго и третьего порядков и постараемся подметить закономерность в образовании этих определителей.

Будем обозначать элементы определителей одной и той же буквой с двумя индексами. Первый индекс будет указывать на номер строки, второй — на номер столбца, занимаемого данным элементом.

При этом определители второго и третьего порядков запишутся в виде:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11},$$

а соответствующие им матрицы — в виде:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \text{ и } \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Определитель второго порядка составлен из элементов матрицы, содержащей 2^2 элементов, а определитель третьего порядка — из элементов матрицы, содержащей 3^2 элементов. Каждый из этих определителей представляет собой алгебраическую сумму членов, причем у определителя второго порядка членов $2=2!$, а у определителя третьего порядка 6, т. е. $3!$.

Каждый член определителя второго порядка есть произведение двух, а каждый член определителя третьего порядка есть произведение трех элементов соответствующей матрицы, причем в любом из этих произведений имеется по одному и только по одному элементу из каждой строки и каждого столбца матрицы¹⁾. Членами каждого из определителей являются все произведения указанного вида, какие только можно составить из элементов соответствующей определителю матрицы.

Для того чтобы установить общее правило расстановки знаков у членов определителей, нам придется несколько отвлечься от определителей и пополнить имеющиеся сведения о перестановках.

Перестановки. Рассмотрим n первых натуральных чисел $1, 2, 3, \dots, n$. Как известно, из этих чисел можно составить $n!$ перестановок²⁾.

Будем говорить, что числа i и j в данной перестановке составляют *инверсию*, если $i > j$, а в перестановке число i стоит раньше, чем число j .

Так, в перестановке $1, 2, 4, 5, 3$ числа 4 и 3, а также 5 и 3 составляют инверсию.

Подсчет числа инверсий в данной перестановке можно выполнять следующим образом: найти, сколько цифр написано впереди единицы, зачеркнуть единицу, затем подсчитать, сколько цифр осталось (не считая зачеркнутой) впереди двух, зачеркнуть цифру «два» и т. д.

Если перед единицей было m_1 цифр, перед двумя — m_2 цифр, ..., перед $n - m_n$ цифр, то общее число инверсий в перестановке равно: $m_1 + m_2 + \dots + m_n$.

Например, в перестановке $1, 3, 5, 4, 2$ число инверсий $t=4$ (перед единицей цифр нет), $\lambda, 3, 5, 4, 2$, перед двумя — три цифры, $\lambda, 3, 5, 4, 2$, перед тремя — нет, $\lambda, 3, 5, 4, 2$, перед четырьмя — одна цифра, $\lambda, 3, 5, 4, 2$, перед пятью — нет, итого $t=0+3+0+1+0=4$.

Перестановка называется *четной*, если число ее инверсий четное, и *нечетной* — в противном случае³⁾. Рассмотренная выше перестановка $1, 3, 5, 4, 2$ — четная.

Вернемся к исследованию определителей. Для установления общего правила расстановки знаков членов запишем у каждого члена множители в порядке возрастания первых индексов.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \quad (1)$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \quad (2)$$

Рассмотрим перестановки, составляемые вторыми индексами элементов членов, и подсчитаем в них число инверсий.

Для определителя второго порядка имеем перестановки $1, 2$ (0 инверсий) и $2, 1$ (1 инверсия). Для определителя третьего порядка — перестановки: $1, 2, 3$ (0 инверсий), $2, 3, 1$ (2 инверсии), $3, 1, 2$ (2 инверсии), $3, 2, 1$ (3 инверсии), $2, 1, 3$ (1 инверсия), $1, 3, 2$ (1 инверсия).

¹⁾ Последнее особенно хорошо видно благодаря введенной выше системе двойных индексов элементов.

²⁾ Напомним, что *перестановкой* из чисел $1, 2, \dots, n$ называется всякое расположение этих чисел в некотором определенном порядке.

³⁾ Число «нуль» считается четным.

Сопоставляя четность числа инверсий в этих перестановках со знаком соответствующего члена, заключаем, что и у определителя (1) (второго порядка) и у определителя (2) (третьего порядка) знак плюс имеют члены, у которых перестановка вторых индексов четная, и знак минус — члены, у которых эта перестановка нечетная. Установленную выше закономерность в составлении определителей второго и третьего порядков можно положить в основу обобщения понятия определителя.

Пусть дана квадратная таблица (*матрица*) из n^2 чисел (*элементов матрицы*):

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \quad (3)$$

Определителем n -го порядка, соответствующим матрице (3), называется число D , равное алгебраической сумме $n!$ членов, каждый из которых представляет собой произведение n элементов, взятых по одному и только по одному из каждой строки и каждого столбца матрицы. Знак члена совпадает со знаком $(-1)^t$, где t — число инверсий в перестановке вторых индексов элементов члена при условии, что первые индексы расположены в порядке возрастания.

Определитель D принято обозначать символом:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Таким образом, по определению, определитель D есть число, равное сумме произведений: $(-1)^t a_{1j_1} a_{2j_2} \dots a_{nj_n}$, где $j_1, j_2, \dots, j_n$ — перестановка из чисел $1, 2, \dots, n$; t — число инверсий в этой перестановке, причем в сумму входят произведения указанного вида со всевозможными перестановками $j_1, j_2, \dots, j_n$ из чисел $1, 2, \dots, n$.

Данное определение включает в себя, как частный случай, введенные ранее определители второго ($n=2$) и третьего ($n=3$) порядков. При $n=1$ матрица (3) состоит из одного элемента a_{11} , соответственно определитель «первого порядка» считается равным числу a_{11} .

Пример 1. Определить, с каким знаком в определитель пятого порядка входит член

$$a_{12} a_{35} a_{41} a_{24} a_{53}.$$

Для определения знака этого члена располагаем в произведении элементы в порядке возрастания первых индексов:

$$a_{12} a_{24} a_{35} a_{41} a_{53}$$

и рассматриваем перестановку вторых индексов: 2, 4, 5, 1, 3.

Так как она содержит $3 + 0 + 2 + 0 + 0 = 5$ инверсий, знак члена совпадает со знаком числа $(-1)^5$.

Пример 2. Вычислить определитель четвертого порядка

$$D = \begin{vmatrix} 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 5 \\ -1 & 2 & 0 & 0 \end{vmatrix}.$$

Запишем определитель D в виде:

$$D = \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ a_{41} & a_{42} & 0 & 0 \end{vmatrix}.$$

Каждый член определителя D есть произведение четырех элементов, взятых по одному из каждой строки и каждого столбца определителя. Всего таких членов будет $4! = 24$. Выписывая ненулевые члены и учитывая их знаки, получим:

$$D = (-1)^2 a_{11}a_{23}a_{34}a_{42} + (-1)^3 a_{12}a_{23}a_{34}a_{41}.$$

Подставляя значения элементов, имеем:

$$D = 2 \cdot 3 \cdot 5 \cdot 2 - 1 \cdot 3 \cdot 5 \cdot (-1) = 75.$$

Для определения знака члена определителя не обязательно располагать элементы члена в порядке возрастания первых индексов. Об этом говорит следующая теорема.

Теорема. Знак члена $a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_n j_n}$ определителя n -го порядка совпадает со знаком $(-1)^{s+t}$, где s — число инверсий в перестановке $i_1, i_2, \dots, i_n$ первых индексов, а t — число инверсий в перестановке $j_1, j_2, \dots, j_n$ вторых индексов.

Доказательство этой теоремы в данном курсе не рассматривается.

На основании этой теоремы в примере 1 для определения знака члена $a_{12}a_{35}a_{41}a_{24}a_{33}$ достаточно было рассмотреть перестановки 1, 3, 4, 2, 5 и 2, 5, 1, 4, 3 первых и вторых индексов элементов и подсчитать в каждой из них число инверсий. Так как $s=2$, $t=5$, то знак члена совпадает со знаком числа $(-1)^7$.

Предлагаемое обобщение понятия определителя на случай любого порядка n позволяет перенести на определители высшего порядка установленные ранее свойства определителей второго и третьего порядков. Не останавливаясь на доказательствах, перечислим эти свойства определителей любого порядка.

1. Определитель не изменится, если его строки заменить столбцами, и наоборот.

2. От перестановки двух столбцов или двух строк определитель меняет только знак.

3. Определитель, имеющий два одинаковых столбца (строки), равен нулю.

4. Если все элементы какого-либо столбца (строки) определителя умножить на одно и то же число m , то определитель умножится на это число [иначе, общий множитель элементов какого-либо столбца (строки) можно вынести за знак определителя].

5. Определитель, у которого элементы двух столбцов (строк) соответственно пропорциональны, равен нулю.

6. Если каждый элемент какого-либо столбца (строки) определителя есть сумма двух слагаемых, то определитель равен сумме двух определителей. У одного из них элементами соответствующего столбца (строки) будут первые слагаемые, у другого — вторые. Остальные элементы у этих двух определителей те же, что и у данного.

7. Определитель не изменится, если к элементам какого-либо столбца (строки) прибавить соответствующие элементы другого столбца (строки), умноженные на одно и то же число.

Если исходить из определения, то для вычисления определителя n -го порядка нужно найти алгебраическую сумму $n!$ членов, что при $n > 3$, вообще говоря, весьма затруднительно (так, если $n=4$, то $n! = 24$). Применение свойств 1—7 позволяет упрощать вычисление определителей.

Условимся в дальнейшем для краткости выражения «умножим элементы столбца (строки) на число m », «прибавим к элементам столбца (строки) элементы другого столбца (строки), умноженные на число m » заменять выражениями: «умножим столбец (строку) на число m », «прибавим к столбцу (строке) другой столбец (строку), умноженный на число m » и т. д.

Пример. Вычислим определитель

$$D = \begin{vmatrix} 1 & 0 & -1 & 1 \\ 1 & 2 & 3 & -1 \\ 3 & 0 & -1 & 1 \\ 0 & 1 & 4 & -1 \end{vmatrix}.$$

Прибавляя к третьему столбцу первый столбец, а из четвертого столбца вычитая первый столбец, получим:

$$D = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 4 & -2 \\ 3 & 0 & 2 & -2 \\ 0 & 1 & 4 & -1 \end{vmatrix}.$$

Вынесем за знак определителя общие множители элементов из третьего и четвертого столбцов.

$$D = -2 \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 2 & 2 \\ 3 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 \end{vmatrix}.$$

Вычитаем из четвертого и третьего столбцов второй столбец. При этом

$$D = -2 \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 \end{vmatrix}.$$

Так как каждый член определителя содержит по одному множителю из каждой его строки и столбца, то, выписывая единственный ненулевой член определителя, мы получим:

$$D = -2 [(-1)^1 1 \cdot 2 \cdot 2 \cdot 1] = 8.$$

Рассмотрим матрицу из n^2 элементов:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

и соответствующий ей определитель n -го порядка:

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Минором любого элемента этого определителя называется определитель $(n-1)$ -го порядка, соответствующий матрице, полученной из данной вычеркиванием строки и столбца, в которых лежит указанный элемент¹⁾. Минор элемента a_{ij} будем обозначать M_{ij} .

Алгебраическим дополнением любого элемента a_{ij} определителя D называется его минор M_{ij} , взятый со знаком числа $(-1)^{i+j}$.

Алгебраическое дополнение элемента a_{ij} будем обозначать A_{ij} . Таким образом, $A_{ij} = (-1)^{i+j} M_{ij}$.

¹⁾ Будем в дальнейшем говорить, что минор элемента есть определитель, полученный из данного определителя вычеркиванием строки и столбца, в которых лежит этот элемент.

Например, у определителя

$$\begin{vmatrix} 1 & 2 & 3 & -1 \\ 4 & 7 & 11 & 0 \\ 0 & -4 & 2 & 3 \\ 5 & 8 & 7 & 2 \end{vmatrix}$$

алгебраическое дополнение элемента $a_{32} = -4$ равно:

$$A_{32} = (-1)^{3+2} \begin{vmatrix} 1 & 3 & -1 \\ 4 & 11 & 0 \\ 5 & 7 & 2 \end{vmatrix} = -25.$$

Для определителей высшего порядка имеет место следующая теорема о разложении определителя по элементам столбцов и строк.

Теорема. Определитель D равен сумме произведений элементов любого его столбца (строки) на их алгебраические дополнения:

$$D = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj}, \quad j = 1, 2, \dots, n,$$

$$D = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}, \quad i = 1, 2, \dots, n.$$

Доказательство этой теоремы читатель найдет в любом курсе высшей алгебры.

Теорема о разложении определителей по элементам строк и столбцов позволяет вычисление определителя n -го порядка свести к вычислению определителей порядка $n-1$; те, в свою очередь, — к вычислению определителей порядка $n-2$ и т. д.

Эта теорема, в сочетании со свойствами определителей, часто используется на практике для вычисления определителей. Для этого обычно, применяя свойства определителей, получают из данного определителя определитель, у которого все элементы какого-либо столбца (строки), кроме, может быть, одного, равны нулю, затем заменяют, на основании указанной теоремы, его вычисление вычислением одного определителя порядка, на единицу меньшего, и т. д.

Пример 1.

$$D = \begin{vmatrix} 2 & 1 & -1 & 0 \\ 3 & 5 & 4 & -1 \\ 0 & 2 & 1 & 3 \\ 1 & -1 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 & 0 \\ -7 & 5 & 9 & -1 \\ -4 & 2 & 3 & 3 \\ 3 & -1 & 2 & 1 \end{vmatrix} = (-1)^{1+2} \begin{vmatrix} -7 & 9 & -1 \\ -4 & 3 & 3 \\ 3 & 2 & 1 \end{vmatrix} =$$

$$= - \begin{vmatrix} -4 & 11 & 0 \\ -13 & -3 & 0 \\ 3 & 2 & 1 \end{vmatrix} = - \begin{vmatrix} -4 & 11 \\ -13 & -3 \end{vmatrix} = -155.$$

Пример 2. Определитель

$$V_n = \begin{vmatrix} 1 & 1 & \dots & 1 \\ k_1 & k_2 & \dots & k_n \\ k_1^2 & k_2^2 & \dots & k_n^2 \\ \dots & \dots & \dots & \dots \\ k_1^{n-1} & k_2^{n-1} & \dots & k_n^{n-1} \end{vmatrix}.$$

называется *определителем Вандермонда*.

Для его вычисления вычтем из каждой строки предыдущую, умноженную на k_1 .

$$V_n = \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & k_2 - k_1 & \dots & k_n - k_1 \\ 0 & (k_2 - k_1)k_2 & \dots & (k_n - k_1)k_n \\ \dots & \dots & \dots & \dots \\ 0 & (k_2 - k_1)k_2^{n-2} & \dots & (k_n - k_1)k_n^{n-2} \end{vmatrix}.$$

Разлагаем V_n по элементам первого столбца:

$$V_n = \begin{vmatrix} k_2 - k_1 & \dots & k_n - k_1 \\ (k_2 - k_1) k_2 & \dots & (k_n - k_1) k_n \\ (k_2 - k_1) k_2^{n-2} & \dots & (k_n - k_1) k_n^{n-2} \end{vmatrix}.$$

Вынося общие множители столбцов за знак определителя, получим

$$V_n = (k_2 - k_1) (k_3 - k_1) \dots (k_n - k_1) V_{n-1},$$

где

$$V_{n-1} = \begin{vmatrix} 1 & \dots & 1 \\ k_2 & \dots & k_n \\ \dots & \dots & \dots \\ k_2^{n-2} & \dots & k_n^{n-2} \end{vmatrix}$$

— определитель Вандермонда порядка $n-1$. Применяя к нему проведенные выше рассуждения, приходим к определителю V_{n-2} и т. д. Окончательно получим:

$$V_n = (k_2 - k_1) (k_3 - k_1) \dots (k_n - k_1) (k_3 - k_2) \dots (k_n - k_2) \dots (k_n - k_{n-1}).$$

§ 37. Некоторые сведения о системах линейных уравнений

Решение многих математических и прикладных задач сводится к решению систем алгебраических уравнений первой степени относительно неизвестных, или, как говорят, *систем линейных уравнений*. При этом число уравнений может отличаться от числа неизвестных. Система из m линейных уравнений с n неизвестными $x_1, x_2, \dots, x_n$ имеет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \quad (1)$$

Числа $a_{11}, a_{12}, \dots, a_{mn}$ называются *коэффициентами* системы (1). Первый индекс указывает на номер уравнения, в которое входит данный коэффициент, а второй индекс — на номер неизвестного, при котором этот коэффициент поставлен. Так, например, a_{32} — коэффициент из третьего уравнения при неизвестном x_2 .

Числа $b_1, b_2, \dots, b_m$ называются *правыми частями*, или *свободными членами*, системы (1).

Система (1) называется *однородной системой линейных уравнений*, если все ее свободные члены равны нулю: $b_1=b_2=\dots=b_m=0$. Если хотя бы один из свободных членов системы не равен нулю, система называется *неоднородной*.

Решением системы (1) называется любая совокупность чисел $x_1, x_2, \dots, x_n$, подстановка которой в уравнения системы обращает их в тождества. Систему, имеющую хотя бы одно решение, будем называть *совместной* (*определенной*, если решение только одно, и *неопределенной*, если решений больше одного). Если система решений не имеет, то говорят, что она несовместна. Очевидно, что всякая однородная система совместна, так как она имеет «нулевое» решение: $x_1=0, x_2=0, \dots, x_n=0$.

Частные случаи систем линейных уравнений ($n=2, 3, m=2, 3$) рассматривались выше.

Для дальнейшего нам будут нужны некоторые сведения о системах линейных уравнений, имеющих одинаковое число уравнений и неизвестных.

Как уже отмечалось выше, всякая такая система имеет «нулевое» решение $\{x_1=0, x_2=0, \dots, x_n=0\}$. Если определитель D системы (4) не равен нулю, то, как следует из сформулированной выше теоремы, это нулевое решение будет единственным¹⁾. Таким образом, для того чтобы система (4) имела ненулевые решения, необходимо, чтобы определитель системы (4) был равен нулю²⁾. Можно показать, что это условие для существования ненулевых решений является также и достаточным.

¹⁾ Его можно получить, например, по формулам Крамера, так как в данном случае $D_1=0, D_2=0, \dots, D_n=0$.

²⁾ Решение $\{x_1, x_2, \dots, x_n\}$ называется *ненулевым*, если хотя бы одно из чисел $x_1, x_2, \dots, x_n$ отлично от нуля.

ГЛАВА II

АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ В ПРОСТРАНСТВЕ

Раздел I

ДЕКАРТОВА ПРЯМОУГОЛЬНАЯ СИСТЕМА КООРДИНАТ В ПРОСТРАНСТВЕ

Декартова прямоугольная система координат в пространстве считается заданной, если даны единичный (масштабный) отрезок, три взаимно перпендикулярные оси, пересекающиеся в одной точке (обозначим ее O и назовем *началом координат*) и указано, которая из этих осей является первой, какая второй и какая третьей. Первая ось называется *осью абсцисс* (осью OX), вторая — *осью ординат* (осью OY), третья — осью аппликат (осью OZ). Эти оси называются *координатными осями*.

Если не оговорено противное, будем считать, что ось абсцисс расположена горизонтально и направлена вперед, ось ординат расположена

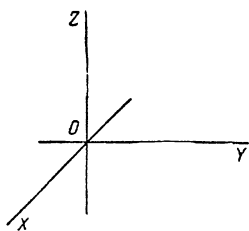


Рис. 68 а.

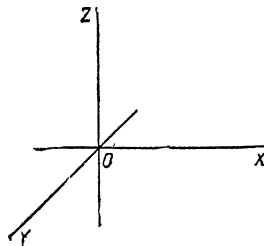


Рис. 68 б.

горизонтально и направлена слева направо, а ось аппликат расположена вертикально и направлена снизу вверх. Положительные направления осей обозначаются стрелками или буквами (соответственно X , Y и Z) (рис. 68, а)

Примечание. Система координат XYZ , изображенная на рис. 68, б, называется *левой*, в отличие от *правой* системы, в которой ось OX направлена вперед, а ось ординат — вправо.

Пусть M — произвольная точка пространства. Обозначим через N , L и K проекции точки M на оси координат (соответственно OX , OY и OZ). Точки N , L и K являются точками пересечения с осями координат плоскостей, перпендикулярных этим осям и проходящих через точку M (рис. 69).

Первой координатой, или абсциссой, точки M называют число x , равное величине отрезка $\overline{ON}$, второй координатой, или ординатой, точки M , называют число y , равное величине отрезка $\overline{OL}$, и, наконец, третьей координатной, или аппликатой, точки M называют число z , равное величине отрезка $\overline{OK}$ (рис. 69).

Если точка M имеет координаты x , y и z , то пишут $M(x, y, z)$ или (x, y, z) .

По определению координаты точки M есть координаты ее проекций на соответствующих осях. Точка O имеет координаты $0, 0, 0$. Любая точка M , лежащая на оси OX , имеет ординату y и аппликату z , равные нулю. Соответственно для точек оси OY $x=0$ и $z=0$, а для точек оси OZ $x=0$ и $y=0$.

Рассмотрим три взаимно перпендикулярные плоскости, проходящие через координатные оси. Плоскость, проходящую через оси OX и OY , обозначим XOY , плоскость, проходящую через оси OY и OZ , обозначим YOZ , а плоскость, проходящую через оси OZ и OX , — ZOX . Эти плоскости называют *координатными плоскостями*.

Для точек, лежащих в плоскости XOY , аппликата z равна нулю. Соответственно для точек плоскости YOZ имеем $x=0$, а для точек плоскости ZOX имеем $y=0$.

Координатные плоскости делят пространство на восемь частей — *октантов*. Первым называется октант, у точек $M(x, y, z)$ которого все координаты положительны. Нумерация остальных октантов приведена на рис. 70.

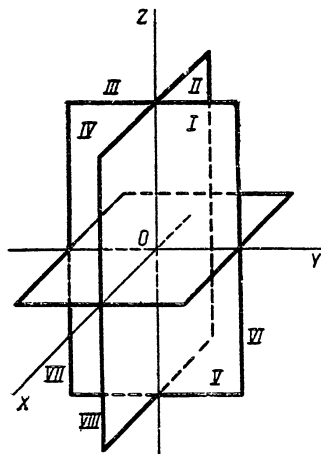


Рис. 70.

лежащая в пересечении этих плоскостей, будет иметь координаты x, y и z .

Вводя в пространстве декартову систему координат, мы получаем возможность определить положение любой точки с помощью троек чисел — ее координат.

Примечание. Для определения положения точки пространства можно, кроме декартовой прямоугольной системы координат, вводить также и другие системы. Некоторые из них (так называемые *цилиндрическая* и *сферическая системы координат*) будут рассматриваться при изучении математического анализа.

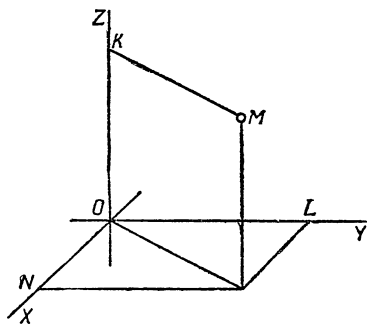


Рис. 69.

Введение в пространстве декартовой прямоугольной системы координат позволяет каждой точке пространства поставить в соответствие упорядоченную тройку действительных чисел — координат этой точки. И наоборот, каждой упорядоченной тройке действительных чисел соответствует в пространстве единственная точка, для которой эти три числа являются ее координатами (в выбранной системе координат).

Для построения точки M по известным ее координатам x, y, z достаточно на оси OX построить точку N , для которой $\overline{ON}=x$, на оси ординат найти точку L , для которой $\overline{OL}=y$, на оси аппликат найти точку K , для которой $\overline{OK}=z$, и через точки N, L и K провести плоскости, перпендикулярные соответственно осям OX, OY и OZ . Точка M ,

ЭЛЕМЕНТЫ ВЕКТОРНОЙ АЛГЕБРЫ

§ 1. Векторные и скалярные величины

Величины, встречающиеся в механике, физике и других прикладных науках, могут быть разделены на две категории. С одной стороны, существуют такие физические или механические величины, которые определяются только числовым значением (числом), например: масса, плотность, температура, объем; с другой — такие величины, которые для своего определения требуют знания не только числового значения, но и направления, например: сила, скорость, ускорение. Величины первой категории называются *скалярными*; величины второй категории называются *векторными*. Скалярная величина может быть задана числом, которое выра-

жает отношение этой величины к соответствующей единице измерения.

Для отвлеченного изображения конкретных векторных величин (физических, механических и т. д.) употребляются векторы. *Векторами* называются направленные отрезки.

Вектор, как направленный отрезок, записывается в виде двух больших латинских букв со стрелкой наверху, причем первая буква обозначает начало вектора, а вторая его конец, напри-

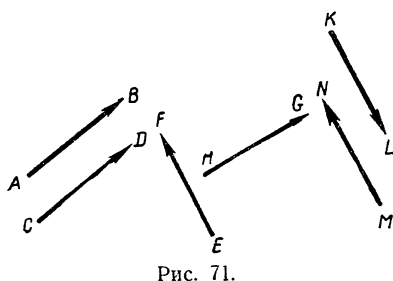


Рис. 71.

мер; $\overrightarrow{MN}$ — вектор, у которого точка M является началом, а N — концом. Иногда вектор обозначают одной малой латинской буквой жирным шрифтом.

Два вектора считаются *равными*, если они параллельны, одинаково направлены и имеют одинаковую длину¹⁾.

Длина MN вектора $\overrightarrow{MN}$ называется его *модулем*. Модуль вектора $\mathbf{a}$ обозначают $|\mathbf{a}|$ или a .

Вектор $\mathbf{0}$, начало и конец которого совпадают, называется *нулевым*. Его модуль равен нулю. Нулевой вектор не имеет определенного направления.

На рис. 71 изображены шесть векторов, равных по модулю, но равными векторами можно назвать лишь $\overrightarrow{AB}$ и $\overrightarrow{CD}$; $\overrightarrow{EF}$ и $\overrightarrow{MN}$:

$$\overrightarrow{AB} = \overrightarrow{CD}; \quad \overrightarrow{EF} = \overrightarrow{MN}.$$

Векторы $\overrightarrow{MN}$ и $\overrightarrow{KL}$ равны по модулю, параллельны, но направлены в противоположные стороны; такие векторы называются *взаимно противоположными*.

¹⁾ В механике, физике и других прикладных науках рассматривают векторы связанные, скользящие и свободные.

Векторы называются *свободными*, если положение их начала не играет никакой роли. Свободные векторы называются *равными*, если они параллельны, имеют одинаковые длины и одинаково направлены.

Скользящие векторы — это такие векторы, которые считаются равными, если они не только имеют одинаковые длины и одинаково направлены, но и расположены на одной и той же прямой.

Связанные векторы — это такие векторы, которые считаются равными, если они не только имеют одинаковые длины и одинаково направлены, но и имеют одинаковое начало.

В дальнейшем изложении будем пользоваться только *с в о б о д н ы м и* векторами.

§ 2. Линейные операции над векторами

Сложение векторов. Суммой векторов a и b , называется такой вектор c , начало которого совпадает с началом вектора a , а конец — с концом вектора b , при условии, что начало вектора b приложено к концу вектора a (рис. 72).

Покажем, что:

1. Сложение векторов подчиняется *переместительному* закону:

$$a + b = b + a \quad (1)$$

для любых векторов a и b .

Приложив векторы a и b к одной точке O , построим на них параллелограмм (рис. 73). Тогда

$$\begin{aligned} a &= \overrightarrow{OK}, & b &= \overrightarrow{OL}, & a + b &= \overrightarrow{OL}, \\ b &= \overrightarrow{OM}, & a &= \overrightarrow{ML}, & b + a &= \overrightarrow{OL}. \end{aligned}$$

Следовательно, $a + b = b + a$.

Упомянутый параллелограмм может вырождаться в отрезок или в точку. Это происходит в случае коллинеарности векторов a и b (два и вообще несколько векторов называются *коллинеарными*, если, будучи приложены к одной точке, они окажутся лежащими на одной прямой). Вырождение в точку произойдет, в частности, в случае нулевых a и b .

2. Сложение векторов подчиняется *сочетательному* закону. Это значит, что для любых векторов a, b, c :

$$(a + b) + c = a + (b + c). \quad (2)$$

Приложив векторы a, b, c к одной точке O , построим на них параллелепипед (рис. 74).

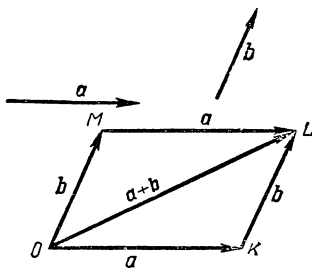


Рис. 73.

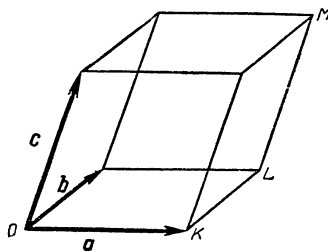


Рис. 74.

Тогда

$$\begin{aligned} a &= \overrightarrow{OK}, & b &= \overrightarrow{KL}, & c &= \overrightarrow{LM}; \\ a + b &= \overrightarrow{OL}, & b + c &= \overrightarrow{KM}, \\ (a + b) + c &= \overrightarrow{OM}; & a + (b + c) &= \overrightarrow{OM} \end{aligned}$$

и, следовательно,

$$(a + b) + c = a + (b + c).$$

В случае *компланарности* векторов (три и вообще несколько векторов называются *компланарными*, если, будучи приложены к одной точке, они окажутся лежащими в одной плоскости) параллелепипед может вырождаться в параллелограмм, в отрезок или в точку. В частности, вырождение произойдет в случае коллинеарности любых двух векторов из векторов a, b, c .

Операция сложения векторов может быть распространена на любое число слагаемых векторов.

Для того чтобы сложить n векторов $a_1, a_2, \dots, a_n$, надо к концу первого вектора приложить начало второго, затем к концу второго вектора приложить начало третьего и т. д. и, наконец, приложить к концу предпоследнего вектора начало последнего; тогда замыкающий вектор, соединяющий начало первого вектора с концом последнего, и будет являться вектором-суммой данных векторов. Из переместительного и сочетательного законов следует, что при сложении векторов можно любым образом переставлять и группировать слагаемые.

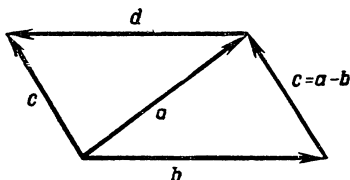


Рис. 75.

Вычитание векторов. Разностью векторов a и b называют вектор c , для которого $c + b = a$.

Непосредственно проверяется, что сумма вектора a и вектора, противоположного b , т. е. $a + (-b)$, есть разность векторов a и b . В самом деле,

$$[a + (-b)] + b = a + [(-b) + b] = a + 0 = a.$$

Этим существование вектора-разности доказано. Остается доказать его единственность. Пусть вектор c таков, что $c + b = a$. Тогда $(c + b) + (-b) = a + (-b)$.

Преобразовывая левую часть равенства получаем:

$$c + [b + (-b)] = c + 0 = c.$$

Этим доказано, что только вектор $a + (-b)$ является разностью векторов a и b . В дальнейшем разность векторов будем записывать в виде $a - b$.

Для геометрического построения вектора-разности c двух векторов можно поступать одним из двух способов.

Первый способ: приложим к некоторому началу вектор-уменьшаемое a и вектор-вычитаемое b (рис. 75) и из конца вектора-вычитаемого b к концу вектора-уменьшаемого проведем вектор c , который и явится искомым вектором-разностью.

Второй способ: приложим к концу вектора a вектор d , противоположный вектору b (т. е. равный вектору $-b$), тогда вектор, соединяющий начало вектора a с концом вектора d , и явится искомым вектором c .

Умножение вектора на скаляр. Произведением вектора a на скаляр λ называется вектор, параллельный a , направленный, как a , если $\lambda \geq 0$, и противоположно, если $\lambda < 0$, и имеющий длину $|\lambda| |a|$. Произведение вектора a на скаляр λ обозначается $a\lambda$ или λa . Если $\lambda = 0$, то по определению $\lambda a = 0$.

Пусть дан некоторый вектор a (рис. 76) и требуется найти произведения $3a$ и $-2a$.

Для получения произведения $3a$ на прямой, на которой расположен вектор a , в направлении вектора a отложим отрезок OM , длина которого равна утроенному модулю вектора a , тогда вектор $\overrightarrow{OM} = 3a$, затем на этой же прямой в направлении, противоположном направлению вектора a , построим вектор $\overrightarrow{ON}$, модуль которого равен двум модулям вектора a , $\overrightarrow{ON} = -2a$.

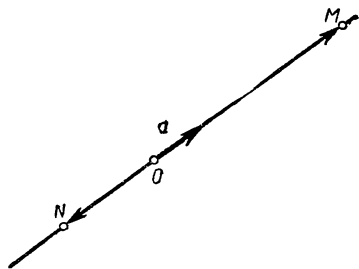


Рис. 76.

Умножение вектора на скаляр подчиняется *сочетательному* закону:

$$\lambda(\mu a) = (\lambda\mu)a. \quad (3)$$

В самом деле, из определения операции умножения вектора на скаляр следует, что при λ и μ , отличных от нуля, векторы $\lambda(\mu a)$ и $(\lambda\mu)a$ параллельны и одинаково направлены (если λ или μ равны 0, то оба вектора нулевые). Так как

$$|\lambda(\mu a)| = |\lambda| |\mu a| = |\lambda| |\mu| |a| = |\lambda\mu| |a|,$$

то эти векторы имеют одинаковую длину.

Умножение вектора на скаляр подчиняется *распределительным* законам:

$$\begin{aligned} (\lambda + \mu)a &= \lambda a + \mu a, \\ \lambda(a + b) &= \lambda a + \lambda b. \end{aligned}$$

В самом деле, первое равенство следует из правила сложения направленных отрезков на прямой, несущей векторы, а второе — из свойств подобных фигур (рис. 77).

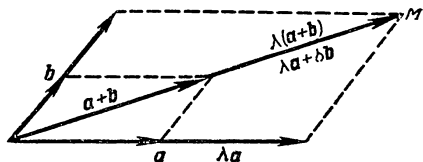


Рис. 77.

§ 3. Основные теоремы о проекциях векторов

Проекцией вектора $\overrightarrow{M_1N}$ на ось называется величина отрезка $\overrightarrow{M_1N_1}$, где M_1 проекция начала вектора, а N_1 проекция конца вектора на эту ось.

Основные теоремы о проекциях векторов формулируются следующим образом.

1. Проекция вектора на какую-либо ось равна произведению модуля вектора на косинус угла наклона вектора к оси¹⁾.

2. Проекция суммы векторов на какую-либо ось равна сумме проекций слагаемых векторов на эту же ось. (Доказательство проводится, как в § 5.)

§ 4. Координаты и компоненты векторов

Координаты вектора. Введем в пространстве декартову прямоугольную систему координат XYZ . Возьмем произвольный вектор a , точку его приложения поместим в начале координат O , а конец его обозначим буквой M (рис. 78). Вектор $\overrightarrow{OM}$ будем называть радиусом-вектором точки M . Пусть координаты точки M будут соответственно: $\overrightarrow{ON}=x$, $\overrightarrow{OL}=y$, $\overrightarrow{OK}=z$.

Спроектируем теперь вектор $a = \overrightarrow{OM}$ на координатные оси. Проекции вектора a на координатные оси называются координатами вектора a . Будем обозначать их X , Y и Z : $\text{пр}_{Ox}a = X$, $\text{пр}_{Oy}a = Y$, $\text{пр}_{Oz}a = Z$. Тот факт, что координаты вектора a равны X , Y , Z , записывается следующим образом:

$$a \{X, Y, Z\}. \quad (1)$$

¹⁾ Угол между векторами, лучами и направленными прямыми в пространстве определяют, как в стереометрии. При радианном измерении он удовлетворяет неравенствам: $0 \leq \varphi \leq \pi$.

Координаты x, y, z точки M (конца вектора a) соответственно равны координатам X, Y, Z вектора a , т. е.

$$X = x, Y = y, Z = z. \quad (2)$$

Разложение векторов на компоненты по ортам в пространстве. Построим тройку векторов i, j, k , удовлетворяющих следующим условиям: 1) начало каждого из них лежит в начале координат; 2) по модулю они равны единице: $|i| = 1; |j| = 1; |k| = 1$; 3) вектор i — направлен по направлению оси абсцисс; j — по направлению оси ординат, k — по направлению оси аппликат.

Эти единичные векторы называются *основными векторами*, или *ортами*. Докажем следующую *теорему*: каждый вектор в пространстве может быть представлен как линейная комбинация ортов (основных векторов).

Линейной комбинацией нескольких векторов называется сумма произведений этих векторов на скаляры.

Возьмем произвольный вектор $a \{X, Y, Z\}$, начало которого поместим в начале координат: $a = \vec{OM}$ (рис. 79). Опустим из точки M на

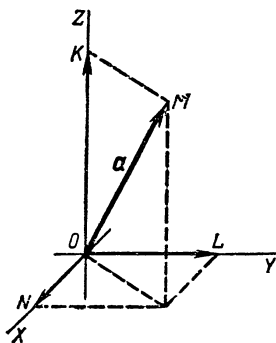


Рис. 78.

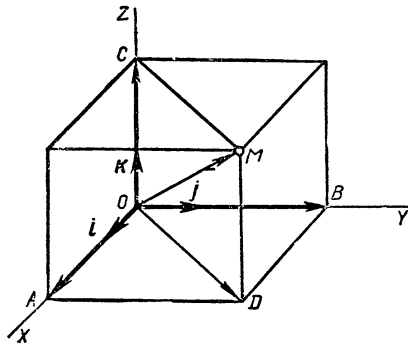


Рис. 79.

плоскость XOY перпендикуляр MD , а затем из точки D опустим перпендикуляры DA на ось OX и DB на ось OY . Опустим также перпендикуляр MC из точки M на ось OZ . Возьмем три вектора $\vec{OA}, \vec{OB}, \vec{OC}$ и сложим их:

$$\begin{aligned} \vec{OA} + \vec{OB} &= \vec{OD}, \\ \vec{OD} + \vec{OC} &= \vec{OM}. \end{aligned}$$

Очевидно, что

$$\vec{OA} + \vec{OB} + \vec{OC} = \vec{OM}. \quad (3)$$

Так как вектор $\vec{OB}$ коллинеарен вектору j , то всегда можно подобрать некоторый множитель μ , такой, что при умножении вектора j на этот множитель будет выполняться равенство

$$\vec{OB} = \mu j. \quad (4)$$

Рассуждая аналогично, можно векторы $\vec{OA}$ и $\vec{OC}$ представить в виде:

$$\vec{OA} = \lambda i \quad (5)$$

$$\vec{OC} = \eta k. \quad (6)$$

Следовательно, равенство (3) можно записать так:

$$a = \lambda i + \mu j + \eta k. \quad (7)$$

Слагаемые $\lambda i, \mu j, \eta k$ называются *компонентами* (или *составляющими*) вектора a . Тройка векторов i, j, k , по которым осуществлено разложение вектора a , называется *координатным базисом*.

Представление вектора в виде суммы компонент называется разложением этого вектора по базису i, j, k . Так как $\overline{OA} = \lambda i$, где i — единичный вектор, то число λ есть отношение $\overline{OA}$ к единице измерения, взятое со знаком плюс или минус в зависимости от того, направлен ли отрезок $\overline{OA}$ так же, как и вектор i , или противоположно ему. Другими словами, λ есть величина отрезка $\overline{OA}$ оси OX , т. е. $\lambda = \overline{OA}$. Но $\overline{OA}$ есть проекция вектора a на ось OX . Следовательно,

$$\lambda = \overline{OA} = \text{пр}_{OX} a = X.$$

Аналогично получаем, что $\mu = Y$, $\eta = Z$, где X, Y, Z — координаты вектора a . При этом равенство (7) запишется следующим образом:

$$a = Xi + Yj + Zk. \quad (8)$$

§ 5. Направляющие косинусы вектора

Модуль вектора. Пусть дан произвольный вектор $a = Xi + Yj + Zk$, требуется определить модуль этого вектора.

Построим в системе координат вектор a и найдем его проекции $X = \overline{OC}$; $Y = \overline{OB}$; $Z = \overline{OD}$ на координатные оси (рис. 80). Беря отрезки OC, OB, OD за ребра параллелепипеда, построим на них прямоугольный параллелепипед. Тогда $OM = |a|$, как диагональ прямоугольного параллелепипеда, можно выразить следующим образом:

$$|a|^2 = X^2 + Y^2 + Z^2$$

или

$$|a| = \sqrt{X^2 + Y^2 + Z^2}. \quad (1)$$

Формула (1) и определяет *модуль (длину)* вектора.

Направляющие косинусы. Обозначим углы, которые образует вектор a с координатными осями, соответственно через α, β, γ (рис. 80).

Так как X, Y, Z являются проекциями вектора на координатные оси, то, следовательно, будут справедливы следующие три равенства:

$$\begin{aligned} X &= |a| \cos \alpha, \\ Y &= |a| \cos \beta, \\ Z &= |a| \cos \gamma, \end{aligned} \quad (2)$$

а так как $|a| = \sqrt{X^2 + Y^2 + Z^2}$, то

$$\begin{aligned} \cos \alpha &= \frac{X}{\sqrt{X^2 + Y^2 + Z^2}}, \\ \cos \beta &= \frac{Y}{\sqrt{X^2 + Y^2 + Z^2}}, \\ \cos \gamma &= \frac{Z}{\sqrt{X^2 + Y^2 + Z^2}}. \end{aligned} \quad (3)$$

При этом $\cos \alpha, \cos \beta, \cos \gamma$ называются *направляющими косинусами* вектора. Покажем, что между углами α, β, γ существует определенная зависимость. Для этого возведем все равенства (3) в квадрат и сложим их. Получим:

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1. \quad (4)$$

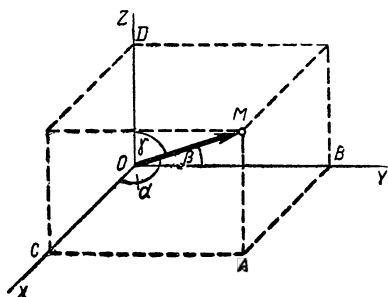


Рис. 80.

§ 6. Лине́йные операции над векторами, заданными в координатной форме

Прежде всего заметим, что проекция суммы векторов на любую ось равна сумме проекций слагаемых векторов на ту же ось. Как следствие из этого предложения и определения координат вектора вытекают следующие правила:

1. Два вектора равны, если равны их координаты.

2. При сложении векторов, заданных в координатной форме, координаты их складываются:

$$a = X_1 i + Y_1 j + Z_1 k; \quad b = X_2 i + Y_2 j + Z_2 k,$$

то

$$(a + b) = (X_1 + X_2) i + (Y_1 + Y_2) j + (Z_1 + Z_2) k. \quad (1)$$

3. При вычитании векторов, заданных в координатной форме, координаты их вычитаются:

$$a = X_1 i + Y_1 j + Z_1 k; \quad b = X_2 i + Y_2 j + Z_2 k,$$

то

$$(a - b) = (X_1 - X_2) i + (Y_1 - Y_2) j + (Z_1 - Z_2) k. \quad (2)$$

4. При умножении вектора, заданного в координатной форме, на скаляр надо все координаты умножить на этот скаляр: если $a = Xi + Yj + Zk$, λ — скалярный множитель, то

$$\lambda a = \lambda Xi + \lambda Yj + \lambda Zk. \quad (3)$$

Коллинеарность двух векторов. Пусть нам даны два ненулевых коллинеарных между собой вектора: $a = X_1 i + Y_1 j + Z_1 k$ и $b = X_2 i + Y_2 j + Z_2 k$.

Если вектор b коллинеарен вектору a , то всегда можно найти такой постоянный множитель λ , что $\lambda b = a$.

В координатной форме имеем:

$$\lambda b = \lambda X_2 i + \lambda Y_2 j + \lambda Z_2 k.$$

Векторы a и b равны, т. е. равны и их координаты, следовательно,

$$X_1 = \lambda X_2; \quad Y_1 = \lambda Y_2; \quad Z_1 = \lambda Z_2. \quad (4)$$

Определяя λ из каждого из трех равенств (4), можно составить пропорции

$$\frac{X_1}{X_2} = \frac{Y_1}{Y_2} = \frac{Z_1}{Z_2} \quad ^1). \quad (5)$$

Наоборот, если координаты векторов a и b удовлетворяют условиям (5), то $\lambda b = a$, т. е. a и b коллинеарны. Равенство (5) является *условием коллинеарности* двух ненулевых векторов. Векторы коллинеарны тогда и только тогда, когда их координаты пропорциональны.

§ 7. Скалярное произведение векторов

В векторной алгебре рассматриваются два вида произведения двух векторов: *скалярное* и *векторное*. Результатом скалярного умножения двух векторов является число (скаляр); результатом векторного умножения

¹⁾ Эти равенства имеют смысл, если знаменатели отношений не равны нулю. В случае, если в каком-либо отношении знаменатель равен нулю, то равен нулю и числитель и соответствующее отношение надо отбросить.

двух векторов является вектор. *Скалярным произведением* двух векторов называется число, равное произведению их длин (модулей), умноженному на косинус угла φ между векторами¹⁾. Для обозначения скалярного произведения вектора $\mathbf{a}$ на вектор $\mathbf{b}$ употребляется одна из записей:

$$(\mathbf{a}, \mathbf{b}), \quad \mathbf{a} \cdot \mathbf{b}, \quad \mathbf{ab}.$$

Согласно определению имеем

$$(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| |\mathbf{b}| \cos \varphi$$

или

$$(\mathbf{a}, \mathbf{b}) = \mathbf{ab} \cos \varphi. \quad (1)$$

Очевидно, что скалярное произведение векторов $\mathbf{a}$ и $\mathbf{b}$ будет равно нулю, если хотя бы один из векторов $\mathbf{a}$ и $\mathbf{b}$ является нулевым вектором или если векторы перпендикулярны (ортогональны). Наоборот, если $(\mathbf{a}, \mathbf{b}) = 0$, то либо хотя бы один из векторов $\mathbf{a}$ и $\mathbf{b}$ — нулевой, либо эти векторы перпендикулярны.

Заметив, что

$$|\mathbf{b}| \cos \varphi = \text{пр}_\mathbf{a} \mathbf{b}, \quad (2)$$

равенство (1) можно записать в виде:

$$(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| \text{пр}_\mathbf{a} \mathbf{b}. \quad (3)$$

Таким образом, скалярное произведение двух векторов равно длине одного вектора, умноженной на проекцию второго на направление первого.

Свойства скалярного произведения

1. Скалярное произведение подчиняется *переместительному закону*, т. е. $(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a})$ для любых векторов $\mathbf{a}$ и $\mathbf{b}$. Действительно,

$$(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| |\mathbf{b}| \cos \varphi, \quad (4)$$

$$(\mathbf{b}, \mathbf{a}) = |\mathbf{b}| |\mathbf{a}| \cos \varphi, \quad (5)$$

т. е. левые части равенств (4) и (5) равны.

2. Скалярное произведение подчиняется *сочетательному закону* относительно скалярного множителя, т. е. $(\lambda \mathbf{a}, \mathbf{b}) = \lambda (\mathbf{a}, \mathbf{b})$ для любых векторов $\mathbf{a}$ и $\mathbf{b}$ и любого числа λ .

Рассмотрим случай, когда $\lambda > 0$:

$$(\lambda \mathbf{a}, \mathbf{b}) = |\mathbf{b}| \text{пр}_\mathbf{b} \lambda \mathbf{a} = |\mathbf{b}| \lambda \text{пр}_\mathbf{b} \mathbf{a} = \lambda |\mathbf{b}| \text{пр}_\mathbf{b} \mathbf{a} = \lambda (\mathbf{a}, \mathbf{b}) \quad (6)$$

или, иначе:

$$(\lambda \mathbf{a}, \mathbf{b}) = |\lambda \mathbf{a}| |\mathbf{b}| \cos \varphi = \lambda |\mathbf{a}| |\mathbf{b}| \cos \varphi = \lambda (\mathbf{a}, \mathbf{b}).$$

Аналогично рассматривается случай, когда $\lambda < 0$.

3. Скалярное произведение подчиняется *распределительному закону*, т. е. для любых векторов $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$

$$\mathbf{a} (\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac}. \quad (7)$$

Действительно,

$$\mathbf{a} (\mathbf{b} + \mathbf{c}) = |\mathbf{a}| \text{пр}_\mathbf{a} (\mathbf{b} + \mathbf{c}) = |\mathbf{a}| (\text{пр}_\mathbf{a} \mathbf{b} + \text{пр}_\mathbf{a} \mathbf{c}) = |\mathbf{a}| \text{пр}_\mathbf{a} \mathbf{b} + |\mathbf{a}| \text{пр}_\mathbf{a} \mathbf{c}. \quad (8)$$

Но так как

$$|\mathbf{a}| \text{пр}_\mathbf{a} \mathbf{b} = (\mathbf{a}, \mathbf{b}) \quad \text{и} \quad \mathbf{a} \text{пр}_\mathbf{a} \mathbf{c} = (\mathbf{a}, \mathbf{c}),$$

то, следовательно,

$$\mathbf{a} (\mathbf{b} + \mathbf{c}) = (\mathbf{a}, \mathbf{b}) + (\mathbf{a}, \mathbf{c}).$$

¹⁾ Если хотя бы один из векторов $\mathbf{a}$ и $\mathbf{b}$ — нулевой, то φ следует считать произвольным.

4. Скалярный квадрат вектора равен квадрату его модуля. Умножим скалярно вектор $\mathbf{a}$ на вектор $\mathbf{a}$:

$$(\mathbf{a}, \mathbf{a}) = |\mathbf{a}| |\mathbf{a}| \cos 0 = |\mathbf{a}|^2. \quad (9)$$

Левую часть равенства (8) принято обозначать $(\mathbf{a})^2$ и называть *скалярным квадратом* вектора $\mathbf{a}$. Тогда равенство (9) примет вид:

$$(\mathbf{a})^2 = |\mathbf{a}|^2. \quad (10)$$

§ 8. Скалярное произведение векторов в координатной форме

Предварительно рассмотрим вопрос о скалярном произведении ортов. Так как разноименные орты перпендикулярны друг другу, а одноименные — равны, то

$$\begin{aligned} (\mathbf{i}, \mathbf{i}) = (\mathbf{j})^2 = |\mathbf{i}|^2 = 1; \quad (\mathbf{j}, \mathbf{j}) = 1; \quad (\mathbf{k}, \mathbf{k}) = 1. \\ (\mathbf{i}, \mathbf{j}) = 0; \quad (\mathbf{j}, \mathbf{k}) = 0; \quad (\mathbf{k}, \mathbf{i}) = 0. \end{aligned} \quad (1)$$

Таким образом, скалярное произведение одноименных ортов равно единице, а разноименных — равно нулю.

Возьмем теперь два вектора, заданных в координатной форме:

$$\mathbf{a} = X_1\mathbf{i} + Y_1\mathbf{j} + Z_1\mathbf{k}; \quad \mathbf{b} = X_2\mathbf{i} + Y_2\mathbf{j} + Z_2\mathbf{k},$$

и перемножим их скалярно. Правые части можно перемножить по правилу умножения многочлена на многочлен, так как скалярное произведение подчиняется распределительному закону:

$$\begin{aligned} (\mathbf{a}, \mathbf{b}) &= (X_1\mathbf{i} + Y_1\mathbf{j} + Z_1\mathbf{k})(X_2\mathbf{i} + Y_2\mathbf{j} + Z_2\mathbf{k}) = \\ &= X_1X_2(\mathbf{i}, \mathbf{i}) + X_1Y_2(\mathbf{i}, \mathbf{j}) + X_1Z_2(\mathbf{i}, \mathbf{k}) + X_2Y_1(\mathbf{j}, \mathbf{i}) + Y_1Y_2(\mathbf{j}, \mathbf{j}) + \\ &\quad + Y_1Z_2(\mathbf{j}, \mathbf{k}) + X_2Z_1(\mathbf{k}, \mathbf{i}) + Y_2Z_1(\mathbf{k}, \mathbf{j}) + Z_1Z_2(\mathbf{k}, \mathbf{k}). \end{aligned} \quad (2)$$

На основании (1) равенство (2) примет следующий вид:

$$(\mathbf{a}, \mathbf{b}) = X_1X_2 + Y_1Y_2 + Z_1Z_2. \quad (3)$$

Следовательно, скалярное произведение двух векторов равно сумме произведений одноименных координат этих векторов.

Приложения скалярного произведения

Задача 1. Найти угол между ненулевыми векторами $\mathbf{a} \{X_1, Y_1, Z_1\}$ и $\mathbf{b} \{X_2, Y_2, Z_2\}$. Возьмем скалярное произведение и определим косинус угла φ между этими векторами:

$$\begin{aligned} (\mathbf{a}, \mathbf{b}) &= |\mathbf{a}| |\mathbf{b}| \cos \varphi, \\ \cos \varphi &= \frac{(\mathbf{a}, \mathbf{b})}{|\mathbf{a}| |\mathbf{b}|}. \end{aligned} \quad (4)$$

Используя формулы (1) (§ 5) и (3), выразим $\cos \varphi$ через координаты векторов:

$$\cos \varphi = \frac{X_1X_2 + Y_1Y_2 + Z_1Z_2}{\sqrt{X_1^2 + Y_1^2 + Z_1^2} \sqrt{X_2^2 + Y_2^2 + Z_2^2}}. \quad (5)$$

Если векторы $\mathbf{a}$ и $\mathbf{b}$ перпендикулярны, то $\cos \varphi = 0$ и, следовательно,

$$X_1X_2 + Y_1Y_2 + Z_1Z_2 = 0, \quad (6)$$

и наоборот, если $X_1X_2 + Y_1Y_2 + Z_1Z_2 = 0$, то векторы $\mathbf{a}$ и $\mathbf{b}$ перпендикулярны.

Равенство (6) является *условием ортогональности* (перпендикулярности) двух векторов.

Задача 2. Найти проекцию вектора $\mathbf{b} \{X_2, Y_2, Z_2\}$ на вектор $\mathbf{a} \{X_1, Y_1, Z_1\}$. Из равенства (3) (§ 7) определяем искомую проекцию, а

затем результат по формулам (1) (§ 5) и (3) запишем в координатной форме:

$$(a, b) = |a| \operatorname{pr}_a b,$$

$$\operatorname{pr}_a b = \frac{(a, b)}{|a|} \quad (7)$$

или

$$\operatorname{pr}_a b = \frac{X_1 X_2 + Y_1 Y_2 + Z_1 Z_2}{\sqrt{X_1^2 + Y_1^2 + Z_1^2}}. \quad (8)$$

Пример 1. Вычислить угол φ между векторами $a \{2, -3, 1\}$ и $b \{5, 2, -4\}$.
Решение. Находим угол по формуле (5):

$$\cos \varphi = \frac{2 \cdot 5 - 3 \cdot 2 - 1 \cdot 4}{\sqrt{4 + 9 + 1} \sqrt{25 + 4 + 16}} = 0, \quad \varphi = \frac{\pi}{2},$$

векторы a и b перпендикулярны.

Пример 2. Найти проекцию вектора $b \{5, -7, 3\}$ на вектор $a \{2, -2, 1\}$.
Решение. По формуле (7) имеем:

$$\operatorname{pr}_a b = \frac{2 \cdot 5 + (-2) \cdot (-7) + 1 \cdot 3}{\sqrt{4 + 4 + 1}} = 9.$$

§ 9. Векторное произведение

Векторным произведением неколлинеарных векторов a и b называется вектор c , длина которого равна площади параллелограмма, построенного на векторах a и b , который перпендикулярен плоскости этого параллелограмма и направлен так, что если смотреть с его конца, то кратчайший поворот от вектора a к вектору b происходит против часовой стрелки (рис. 81). В случае, если векторы a и b коллинеарны, их векторным произведением называется нулевой вектор.

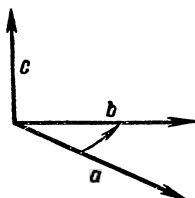


Рис. 81.

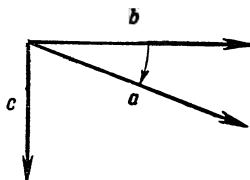


Рис. 82.

Векторное произведение вектора a на вектор b обозначается $a \times b$ или $[a, b]$.

Из определения следует, что если

$$c = [a, b], \quad (1)$$

то

$$|c| = |a| |b| \sin \varphi, \quad (2)$$

где φ — угол между векторами a и b ¹⁾.

Если векторно умножить вектор b на вектор a , то получим вектор $[b, a]$ (рис. 82), равный по модулю вектору $[a, b]$, но направленный в противоположную сторону.

Другими словами, вектор

$$[a, b] = -[b, a]. \quad (3)$$

Из равенства (2) вытекает, что векторное произведение векторов a и b есть нулевой вектор в том и только в том случае, если эти векторы

¹⁾ См. сноску на стр. 103.

коллинеарны. Обращение векторного произведения в нулевой вектор есть условие коллинеарности двух векторов.

Свойства векторного произведения.

1) Для любых векторов a и b

$$[a, b] = -[b, a].$$

Следовательно, векторное произведение *не подчиняется переместительному закону*.

2) Векторное произведение подчиняется *сочетательному закону* относительно скалярного множителя, т. е.

$$[\lambda a, b] = \lambda [a, b], \quad (4)$$

$$[a, \lambda b] = \lambda [a, b] \quad (5)$$

для любых векторов a и b и любого числа λ .

Для доказательства этого свойства достаточно доказать равенство (4). Справедливость равенства (5) устанавливается аналогичным образом.

Будем доказывать равенство (4), взяв $\lambda > 0$.

Пусть $[\lambda a, b] = c$, а $\lambda [a, b] = d$.

Докажем, что векторы c и d равны между собой.

Так как $\lambda > 0$, то направления векторов a и λa совпадают. Следовательно, векторы c и d параллельны и имеют одинаковое направление.

$$|c| = |[\lambda a, b]| = |\lambda a| |b| \sin \varphi = \lambda |a| |b| \sin \varphi,$$

$$|d| = |\lambda [a, b]| = \lambda |a| |b| \sin \varphi,$$

где φ — угол между векторами a и b (λa и b), т. е. $|c| = |d|$. Итак, $c = d$, т. е. $[\lambda a, b] = \lambda [a, b]$.

Таким же путем можно доказать равенство (4) при $\lambda < 0$.

3) Векторное произведение подчиняется *распределительному закону*:

$$[a, (b + c)] = [a, b] + [a, c] \quad (6)$$

для любых векторов a , b и c .

Доказательство этого свойства векторного произведения приводится в § 11¹⁾.

§ 10. Векторное произведение векторов в координатной форме

Рассмотрим сначала векторные произведения ортов. При векторном умножении одноименных ортов, как коллинеарных, очевидно, в результате будет получаться нулевой вектор:

$$[i, i] = 0; \quad [j, j] = 0; \quad [k, k] = 0. \quad (1)$$

При векторном умножении вектора i на вектор j получится вектор, расположенный перпендикулярно плоскости XOY и направленный по направлению оси OZ . Модуль его будет равен единице, так как

$$|[i, j]| = |i| |j| \sin \frac{\pi}{2} = 1.$$

Следовательно, в результате умножения получим единичный вектор, направленный по направлению оси OZ , т. е. вектор k . Если $[i, j] = k$, то на основании равенства (3) § 9

$$[j, i] = -k.$$

¹⁾ С прямым доказательством распределительного свойства векторного произведения можно ознакомиться в любом полном курсе аналитической геометрии.

Аналогично производится умножение других разноименных ортов:

$$\begin{aligned} [j, k] &= i; & [k, j] &= -i; \\ [k, i] &= j; & [i, k] &= -j. \end{aligned} \quad (2)$$

Возьмем теперь два вектора $a = X_1 i + Y_1 j + Z_1 k$; $b = X_2 i + Y_2 j + Z_2 k$ и рассмотрим их векторное произведение.

Так как векторное произведение подчиняется распределительному закону, то векторы можно умножать по правилу умножения многочлена на многочлен, строго соблюдая порядок расположения множителей. При умножении скалярные множители можно выносить за знак векторного произведения.

$$\begin{aligned} [a, b] &= [X_1 i + Y_1 j + Z_1 k, X_2 i + Y_2 j + Z_2 k] = \\ &= X_1 X_2 [i, i] + X_1 Y_2 [i, j] + X_1 Z_2 [i, k] + Y_1 X_2 [j, i] + \\ &+ Y_1 Y_2 [j, j] + Y_1 Z_2 [j, k] + Z_1 X_2 [k, i] + Z_1 Y_2 [k, j] + \\ &+ Z_1 Z_2 [k, k]. \end{aligned} \quad (3)$$

Используя равенства (1) и (2), упростим, а затем преобразуем правую часть равенства (3):

$$\begin{aligned} [a, b] &= X_1 Y_2 k - X_1 Z_2 j - Y_1 X_2 k + Y_1 Z_2 i + Z_1 X_2 j - Z_1 Y_2 i = \\ &= i(Y_1 Z_2 - Z_1 Y_2) + j(Z_1 X_2 - X_1 Z_2) + k(X_1 Y_2 - Y_1 X_2). \end{aligned} \quad (4)$$

Равенство (4) можно записать с помощью определителей:

$$[a, b] = i \begin{vmatrix} Y_1 Z_1 \\ Y_2 Z_2 \end{vmatrix} + j \begin{vmatrix} Z_1 X_1 \\ Z_2 X_2 \end{vmatrix} + k \begin{vmatrix} X_1 Y_1 \\ X_2 Y_2 \end{vmatrix} \quad (4')$$

или условно в виде

$$[a, b] = \begin{vmatrix} i & j & k \\ X_1 Y_1 Z_1 \\ X_2 Y_2 Z_2 \end{vmatrix}^1. \quad (5)$$

Равенство (5) дает выражение векторного произведения через координаты векторов.

Пример 1. Вычислить площадь S треугольника, построенного на векторах $a \{2, -3, 4\}$ и $b \{1, -2, 5\}$.

Решение. Найдем сначала векторное произведение векторов a и b , и вычислим его модуль. Этот модуль численно равен площади параллелограмма, построенного на векторах a и b . Площадь S данного треугольника равна половине площади этого параллелограмма.

$$\begin{aligned} [a, b] &= \begin{vmatrix} i & j & k \\ 2 & -3 & 4 \\ 1 & -2 & 5 \end{vmatrix} = -7i - 6j - k, \\ |[a, b]| &= \sqrt{49 + 36 + 1} = \sqrt{86}, \\ S &= \frac{1}{2} \sqrt{86}. \end{aligned}$$

Пример 2. С помощью векторного произведения найти синус угла φ между векторами $a \{1, 2, 2\}$ и $b \{2, 3, 1\}$, а затем с помощью скалярного произведения проверить ответ.

Решение. На основании формулы (2) (§ 9):

$$\sin \varphi = \frac{|[a, b]|}{|a| |b|}.$$

¹⁾ Если разложить этот «определитель» по элементам первой строки, мы получим равенство (4').

При этом

$$[a, b] = \begin{vmatrix} i & j & k \\ 1 & 2 & 2 \\ 2 & 3 & 1 \end{vmatrix} = -4i + 3j - k,$$

$$|[a, b]| = \sqrt{16 + 9 + 1} = \sqrt{26},$$

$$\sin \varphi = \frac{\sqrt{26}}{3\sqrt{14}}.$$

По формуле (6) (§ 8) найдем

$$\cos \varphi = \frac{2 + 6 + 2}{3\sqrt{14}} = \frac{10}{3\sqrt{14}}.$$

Если задача сделана правильно, то найденные значения $\sin \varphi$ и $\cos \varphi$ должны удовлетворять основному тригонометрическому тождеству $\sin^2 \varphi + \cos^2 \varphi = 1$.

Проверка показывает, что это тождество выполняется.

§ 11. Смешанное произведение трех векторов

При последовательном умножении векторов a , b , c возможны следующие три случая:

1) Сначала два вектора перемножаются скалярно $(a, b) = \lambda$. Последующее умножение будет являться уже знакомым нам действием умножения вектора на скаляр: λc . Этот случай не представит ничего нового.

2) Сначала два вектора перемножаются векторно $[a, b] = d$, затем полученный вектор d еще раз векторно умножается на вектор c . Такое умножение называется *двойным векторным произведением* и обозначается следующим образом: $[[a, b], c]$. В конце настоящего раздела мы вернемся к двойному векторному произведению.

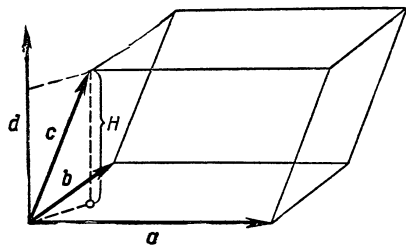


Рис. 83.

3) Два вектора умножаются векторно: $[a, b] = d$, а полученный вектор d скалярно умножается на вектор c . Такое умножение векторов называется *векторно-скалярным*, или *смешанным произведением* трех векторов, и обозначается так:

$$([a, b], c).$$

В результате скалярного умножения (d, c) получим некоторое число. Покажем, что, если векторы a , b , c некопланарны, то это число равно объему параллелепипеда, построенного на векторах a , b , c , взятому со знаком плюс, если взаимное расположение векторов a , b , c совпадает с взаимным расположением осей координат (OX , OY и OZ)¹⁾ и со знаком минус в противном случае.

Возьмем три некопланарных вектора a , b , c , взаимное расположение которых совпадает с взаимным расположением осей координат, и на них как на ребрах построим параллелепипед (рис. 83). Так как $d = [a, b]$, то

$$|d| = |a| |b| \sin \varphi = S, \quad (1)$$

где S — площадь параллелограмма, построенного на векторах a и b . Теперь найдем скалярное произведение (d, c) :

$$(d, c) = |d| \text{ пр. } c, \quad (2)$$

¹⁾ В случае правой системы координат это означает, что если смотреть из конца вектора c , то кратчайший поворот от вектора a к вектору b происходит против часовой стрелки; при этом предполагается, что векторы a , b и c имеют общее начало.

проекция вектора c на d равняется высоте параллелепипеда H (рис. 83). Следовательно,

$$(c, d) = |d| \cdot H. \quad (3)$$

Отсюда

$$([a, b], c) = (d, c) = |d| \cdot H = S \cdot H = V,$$

где V — объем параллелепипеда.

В общем случае

$$([a, b], c) = \pm V. \quad (4)$$

Знак правой части равенства (4), как было указано, определяется взаимным расположением данных векторов.

Из определения смешанного произведения вытекает, что оно обращается в нуль в том и только в том случае, когда векторы компланарны.

Смешанное произведение векторов в координатной форме. Пусть даются три вектора в координатной форме:

$$a = X_1 i + Y_1 j + Z_1 k; \quad b = X_2 i + Y_2 j + Z_2 k; \quad c = X_3 i + Y_3 j + Z_3 k.$$

Требуется найти выражение смешанного произведения $([a, b], c)$.

Запишем сначала векторное произведение векторов a и b в координатной форме:

$$[a, b] = \begin{vmatrix} i & j & k \\ X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{vmatrix} = i \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix} + j \begin{vmatrix} Z_1 & X_1 \\ Z_2 & X_2 \end{vmatrix} + k \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix}. \quad (5)$$

Теперь, используя формулу (3) § 8, найдем скалярное произведение векторов $[a, b]$ и c :

$$([a, b], c) = X_3 \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix} + Y_3 \begin{vmatrix} Z_1 & X_1 \\ Z_2 & X_2 \end{vmatrix} + Z_3 \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \quad (6)$$

или

$$([a, b], c) = \begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}. \quad (7)$$

Формула (7) дает выражение смешанного произведения в координатной форме.

Свойства смешанного произведения. Смешанное произведение трех векторов:

$$a \{X_1, Y_1, Z_1\}; \quad b \{X_2, Y_2, Z_2\}; \quad c \{X_3, Y_3, Z_3\} \quad (8)$$

подчиняется *сочетательному закону*:

$$([a, b], c) = (a, [b, c]) \quad (9)$$

для любых векторов a, b, c .

Действительно, на основании свойства переместительности скалярного произведения имеем:

$$(a, [b, c]) = ([b, c], a). \quad (10)$$

Запишем правую часть равенства (10) в координатной форме, и дважды поменяем местами строки определителя:

$$([b, c], a) = \begin{vmatrix} X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \\ X_1 & Y_1 & Z_1 \end{vmatrix} = - \begin{vmatrix} X_2 & Y_2 & Z_2 \\ X_1 & Y_1 & Z_1 \\ X_3 & Y_3 & Z_3 \end{vmatrix} = \begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}. \quad (11)$$

Сопоставляя правые части равенств (7) и (11), убеждаемся, что

$$([a, b], c) = (a, [b, c]).$$

В связи с этим смешанное произведение векторов a, b, c принято обозначать так: abc .

При круговой¹⁾ перестановке векторов-множителей смешанное произведение не меняется, а при любой другой перестановке меняет знак:

$$abc = cab = bca = -bac = -cba = -acb. \quad (12)$$

Это следует из предыдущего предложения и из правила перестановки множителей в скалярном и векторном произведении двух векторов.

Условие компланарности трех векторов. Как уже отмечалось, смешанное произведение трех компланарных векторов $a \{X_1, Y_1, Z_1\}$; $b \{X_2, Y_2, Z_2\}$; $c \{X_3, Y_3, Z_3\}$ равно нулю:

$$abc = 0 \quad (13)$$

и наоборот, если смешанное произведение abc равно нулю, то векторы a, b, c компланарны. Таким образом, равенство

$$abc = 0$$

есть *условие компланарности* векторов a, b, c .

Используя формулу (7), можно условие компланарности трех векторов записать в следующем виде:

$$\begin{vmatrix} X_1 Y_1 Z_1 \\ X_2 Y_2 Z_2 \\ X_3 Y_3 Z_3 \end{vmatrix} = 0. \quad (14)$$

Пользуясь сочетательным свойством смешанного произведения и распределительным свойством скалярного произведения, легко доказать распределительное свойство векторного произведения.

Прежде всего заметим, что если векторы a и b таковы, что при всяком векторе d имеем $ad = bd$, то $a = b$.

В самом деле, из равенства $ad = bd$ следует $ad - bd = 0$ или, в силу распределительного свойства скалярного произведения,

$$(a - b)d = 0.$$

Беря $d = a - b$, получим $(a - b)^2 = 0$, но скалярный квадрат вектора равен нулю, только если вектор нулевой. Следовательно, $a - b = 0$, т. е. $a = b$.

В силу сделанного замечания для доказательства распределительного закона векторного произведения

$$[(a + b), c] = [a, c] + [b, c]$$

достаточно лишь проверить, что для любого вектора d выполняется равенство

$$[(a + b), c]d = ([a, c] + [b, c])d.$$

Имеем:

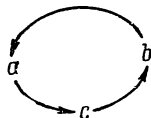
$$\begin{aligned} [(a + b), c]d &= (a + b)[c, d] = a[c, d] + b[c, d] = \\ &= [a, c]d + [b, c]d = ([a, c] + [b, c])d, \end{aligned}$$

что и требовалось доказать.

Двойное векторное произведение. Приведем, без вывода, следующее выражение двойного векторного произведения:

$$[a, [b, c]] = b(a, c) - c(a, b). \quad (15)$$

¹⁾ Круговой перестановкой векторов a, b и c называется такая перестановка, когда b ставится на место a , a — на место c , c — на место b .



Доказательство этого равенства читатель найдет в более полных курсах.

Пример 1. Вычислить объем пирамиды, построенной на трех векторах:

$$a \{2, 3, -1\}, \quad b \{4, 5, 1\} \quad c \{1, -2, 3\}.$$

Решение. Найдем смешанное произведение данных векторов; его модуль равен объему параллелепипеда, построенного на данных векторах. Искомый объем пирамиды представляет одну шестую часть объема этого параллелепипеда.

$$abc = \begin{vmatrix} 2 & 3 & -1 \\ 4 & 5 & 1 \\ 1 & -2 & 3 \end{vmatrix} = 14; \quad V_{\text{пир}} = \frac{14}{6} = \frac{7}{3} e d^3.$$

Пример 2. Проверить, будут ли три данных вектора $a \{2, 1, 3\}$ $b \{3, -2, 1\}$, $c \{1, 1, 2\}$ компланарны.

Решение. Для этого достаточно проверить, выполняется ли условие (14) компланарности векторов:

$$\begin{vmatrix} 2 & 1 & 3 \\ 3 & -2 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 2(-4-1) - (6-1) + 3(3+2) = -10-5+15=0.$$

Раздел 3

ПРОСТЕЙШИЕ ЗАДАЧИ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ В ПРОСТРАНСТВЕ

§ 12. Расстояние между двумя точками в пространстве

В пространстве даны две точки $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$. Требуется определить расстояние между ними.

Проведем радиусы-векторы r_1 и r_2 точек M_1 и M_2 (рис. 84).

Пусть $r_1 \{x_1, y_1, z_1\}$ и $r_2 \{x_2, y_2, z_2\}$. Тогда $\overline{M_1M_2} = r_2 - r_1$, а координатами этого вектора будут $x_2 - x_1, y_2 - y_1, z_2 - z_1$. Следовательно,

$$|r_2 - r_1| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \quad (1)$$

Обозначая расстояние между M_1 и M_2 через d , получим:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \quad (2)$$

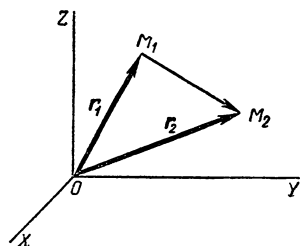


Рис. 84.

§ 13. Деление отрезка в данном отношении

Пусть даны две точки $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$ (рис. 85). Установим на прямой M_1M_2 направление (например, от M_1 к M_2). Требуется найти на этой прямой отрезок $\overline{M_0M_2}$ (x_0, y_0, z_0), которая разделила бы отрезок $\overline{M_1M_2}$ в отношении λ , т. е. так, что $\frac{\overline{M_1M_0}}{\overline{M_0M_2}} = \lambda$. Последнее условие можно записать в форме:

$$\overline{M_1M_0} = \lambda \overline{M_0M_2},$$

откуда следует, что

$$\overrightarrow{M_1M_0} = \lambda \overrightarrow{M_0M_2}. \quad (1)$$

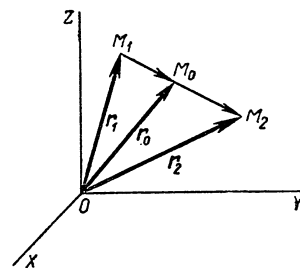


Рис. 85.

Координаты векторов $\overrightarrow{M_1M_0} = \mathbf{r}_0 - \mathbf{r}_1$ и $\lambda \overrightarrow{M_0M_2} = \lambda (\mathbf{r}_2 - \mathbf{r}_0)$ соответственно равны:

$$x_0 - x_1, \quad y_0 - y_1, \quad z_0 - z_1 \quad \text{и} \quad \lambda (x_2 - x_0), \quad \lambda (y_2 - y_0), \quad \lambda (z_2 - z_0).$$

Так как у равных векторов координаты равны, то

$$\begin{aligned} x_0 - x_1 &= \lambda (x_2 - x_0), \\ y_0 - y_1 &= \lambda (y_2 - y_0), \\ z_0 - z_1 &= \lambda (z_2 - z_0). \end{aligned} \quad (2)$$

Определяя x_0 из первого из равенств (2), получим

$$x_0 = \frac{x_1 + \lambda x_2}{1 + \lambda}. \quad (3)$$

Аналогичные выражения получим для y_0 и z_0 . Следовательно,

$$x_0 = \frac{x_1 + \lambda x_2}{1 + \lambda}; \quad y_0 = \frac{y_1 + \lambda y_2}{1 + \lambda}; \quad z_0 = \frac{z_1 + \lambda z_2}{1 + \lambda}. \quad (4)$$

Формулы (4) и являются *формулами деления отрезка в данном отношении*.

В случае $\lambda = 1$ получим формулы деления отрезка пополам:

$$x_0 = \frac{x_1 + x_2}{2}; \quad y_0 = \frac{y_1 + y_2}{2}; \quad z_0 = \frac{z_1 + z_2}{2}. \quad (5)$$

Раздел 4

ПОВЕРХНОСТИ И ЛИНИИ В ПРОСТРАНСТВЕ И ИХ УРАВНЕНИЯ

1. Уравнения с тремя неизвестными. Будем рассматривать уравнения вида $F(x, y, z) = 0$, где $F(x, y, z)$ — какие-нибудь выражения, содержащие x, y и z .

Если выражение $F(x, y, z)$ есть сумма конечного числа слагаемых вида $Ax^l y^m z^k$, где l, m и k — целые неотрицательные, а A — действительные числа, то уравнение называется *алгебраическим уравнением*, а большее из чисел $(l + m + k)$ называется его *степенью*.

Алгебраическое уравнение первой степени относительно x, y и z имеет вид: $Ax + By + Cz + D = 0$ (A, B, C, D — числа), алгебраическое уравнение второй степени относительно x, y, z запишется в виде:

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Hyz + Kx + Ly + Mz + N = 0$$

и т. д.

2. Поверхность и ее уравнение. В аналитической геометрии поверхность рассматривается как геометрическое место точек, ее составляющих.

Введем в пространстве прямоугольную декартову систему координат XYZ и рассмотрим какую-нибудь поверхность. Пусть M — произвольная (или, как говорят, «текущая») точка этой поверхности. Положение точки M должно подчиняться определенному условию, характеризующему данное геометрическое место точек (данную поверхность). Так как положение точки M определяется ее координатами x, y, z , то условие, которому должно удовлетворять положение точки M , сводится к условию, которому должны удовлетворять координаты точки M , т. е. сводится к некоторому уравнению

$$F(x, y, z) = 0. \quad (1)$$

Это уравнение $F(x, y, z) = 0$ называется *уравнением данной поверхности*, если координаты любой точки поверхности ему удовлетворяют, а координаты точек, не лежащих на поверхности, не удовлетворяют.

Таким образом, уравнением поверхности называется соотношение $F(x, y, z) = 0$, которому удовлетворяют координаты точек поверхности (и только они)¹⁾.

Рассмотрим пример на составление уравнения поверхности, заданной как геометрическое место точек. Сферой радиуса R с центром в точке O' является геометрическое место всех точек пространства, находящихся от точки O' на расстоянии R .

Пусть в данной декартовой прямоугольной системе координат точка O' (центр сферы) имеет координаты a, b и c . Обозначим через $M(x, y, z)$ произвольную точку сферы. Записывая условие, определяющее точки сферы, получим: $O'M = R$, т. е.

$$\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} = R,$$

т. е.

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = R^2. \quad (2)$$

Координаты любой точки M сферы удовлетворяют уравнению (2).

Если точка $M(x, y, z)$ не лежит на данной сфере, то $O'M > R$ или $O'M < R$, а значит,

$$(x-a)^2 + (y-b)^2 + (z-c)^2 > R^2$$

или

$$(x-a)^2 + (y-b)^2 + (z-c)^2 < R^2.$$

Таким образом, уравнение (2) является уравнением сферы радиуса R с центром в точке $O'(a, b, c)$.

Поверхности, имеющие в декартовой системе координат алгебраические уравнения первой и второй степени, будем называть *алгебраическими* поверхностями соответственно *первого* и *второго порядков*.

3. Уравнения и соответствующие им поверхности. Пусть дано некоторое уравнение относительно переменных x, y, z :

$$F(x, y, z) = 0. \quad (3)$$

Рассмотрим в пространстве XYZ все точки (x, y, z) , координаты которых удовлетворяют уравнению (3). Геометрическое место точек пространства, координаты которых удовлетворяют уравнению $F(x, y, z) = 0$, будем называть поверхностью, определяемой в выбранной системе координат этим уравнением. При этом само уравнение $F(x, y, z) = 0$ будет уравнением такой поверхности. Это определение поверхности позволяет геометрически иллюстрировать различные уравнения, связывающие переменные x, y и z , в связи с чем оно часто используется при изучении вопросов алгебры и анализа.

Ниже будут рассматриваться различные задачи, связанные с заданием поверхности ее уравнением. Условимся в дальнейшем, вместо того чтобы говорить «дана поверхность, определяемая уравнением $F(x, y, z) = 0$ », говорить короче: «дана поверхность $F(x, y, z) = 0$ » и т. д.

Заметим, что данное выше «определение» поверхности может привести к геометрическим образам, весьма отличным от интуитивного представления о поверхности.

4. Уравнения линии в пространстве. Линию в пространстве можно рассматривать как линию пересечения двух поверхностей, т. е. как геометрическое место точек, общих двум поверхностям²⁾.

Пусть $F(x, y, z) = 0$ и $\Phi(x, y, z) = 0$ — уравнения поверхностей, пересекающихся по данной линии L . Тогда координаты каждой точки линии L удовлетворяют обоим уравнениям.

¹⁾ Строгое определение поверхности дается в специальной математической литературе.

²⁾ Строгое определение линии дается в специальной математической литературе.

Наоборот, каждая точка, координаты которой удовлетворяют обоим уравнениям, лежит на линии пересечения этих поверхностей.

Пару уравнений
$$\begin{cases} F(x, y, z) = 0, \\ \Phi(x, y, z) = 0 \end{cases}$$

будем называть *уравнениями линии* L в пространстве.

Примечания: 1. Так как через линию в пространстве можно проводить различные поверхности, то можно в качестве уравнений линии брать уравнения любых двух поверхностей, пересекающихся по данной линии.

2. Определяя линию как геометрическое место точек (x, y, z) (пространства XYZ), координаты которых удовлетворяют уравнениям

$$\begin{cases} F(x, y, z) = 0, \\ \Phi(x, y, z) = 0 \end{cases}$$

мы можем, кроме геометрических мест точек, которые мы интуитивно понимаем под линиями, получить мнимое геометрическое место точек [поверхности $F(x, y, z) = 0$ и $\Phi(x, y, z) = 0$ не пересекаются], одну, несколько или бесчисленное множество отдельных точек, и т. д.

Раздел 5 ПЛОСКОСТЬ

§ 14. Нормальное уравнение плоскости

Возьмем в пространстве XYZ некоторую плоскость P . Через начало координат проведем прямую (нормаль), перпендикулярную плоскости P (K — точка пересечения этой прямой с плоскостью P) и установим на этой прямой направление от начала O в сторону плоскости P (рис. 86).

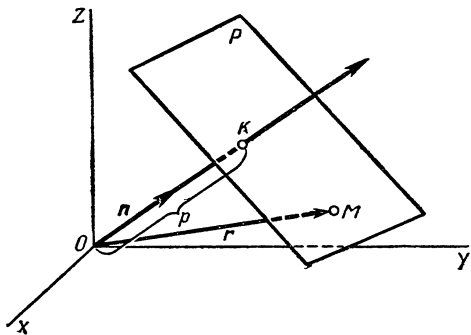


Рис. 86.

Примечание. Если плоскость P проходит через начало O , то на нормали устанавливаем любое из двух возможных направлений.

Пусть n — единичный вектор нормали, направленный в выбранную на ней сторону, α , β и γ — углы наклона этого вектора к осям координат, а p — расстояние плоскости P от начала координат. При этом

$$n \{ \cos \alpha, \cos \beta, \cos \gamma \}, \text{ а } p = OK.$$

Затем возьмем на плоскости P произвольную («текущую») точку $M(x, y, z)$ и соединим ее с началом координат. Радиус-вектор r точки M будет иметь своими координатами числа x, y, z :

$$r = \overrightarrow{OM} \{ x, y, z \}. \quad (1)$$

Спроектируем вектор r на вектор n . Очевидно, что

$$\text{пр}_n r = p. \quad (2)$$

Равенство (2) будет выполняться для всех точек, лежащих на плоскости P .

Так как равенство (2) справедливо только для точек плоскости P , то оно является уравнением этой плоскости. Так как $(\mathbf{r}, \mathbf{n}) = |\mathbf{n}| \text{пр}_{\mathbf{n}} \mathbf{r} = \text{пр}_{\mathbf{n}} \mathbf{r}$, то уравнение плоскости P можно записать в виде:

$$(\mathbf{r}, \mathbf{n}) = p$$

или

$$(\mathbf{r}, \mathbf{n}) - p = 0. \quad (3)$$

Уравнение (3) называется *нормальным уравнением плоскости в векторной форме*. Зная координаты векторов $\mathbf{r}$, $\mathbf{n}$ и выражение для скалярного произведения, можно записать уравнение (3) в координатной форме:

$$x \cos \alpha + y \cos \beta + z \cos \gamma - p = 0. \quad (4)$$

Уравнение (4) называется *нормальным уравнением плоскости в координатной форме*. Параметрами этого уравнения являются: p — расстояние плоскости от начала координат; α , β , γ — углы, которые образует нормаль $\mathbf{n}$ с координатными осями.

Если плоскость P проходит через начало координат, то уравнение плоскости запишется в виде: $(\mathbf{r}, \mathbf{n}) = 0$, т. е.

$$x \cos \alpha + y \cos \beta + z \cos \gamma = 0.$$

Таким образом, доказано, что всякая плоскость в прямоугольной системе координат XYZ имеет уравнение первой степени относительно x , y и z .

§ 15. Общее уравнение плоскости

Возьмем произвольное уравнение первой степени относительно x , y и z :

$$Ax + By + Cz + D = 0^1) \quad (1)$$

и докажем, что оно является уравнением некоторой плоскости.

Обозначим через $\mathbf{N}$ вектор с координатами A , B , C :

$$\mathbf{N} \{A, B, C\}, \quad (2)$$

а через $\mathbf{r}$ — радиус-вектор точки $M(x, y, z)$, координаты которой удовлетворяют уравнению (1). При этом

$$\mathbf{r} \{x, y, z\}. \quad (3)$$

Так как скалярное произведение векторов $\mathbf{N}$ и $\mathbf{r}$ равно

$$(\mathbf{N}, \mathbf{r}) = Ax + By + Cz, \quad (4)$$

то уравнение (1) можно переписать в виде:

$$(\mathbf{N}, \mathbf{r}) + D = 0. \quad (5)$$

Умножая все члены уравнения (5) на некоторый множитель λ , получим уравнение, являющееся уравнением того же геометрического образа:

$$(\lambda \mathbf{N}, \mathbf{r}) + \lambda D = 0. \quad (6)$$

Выберем теперь число λ так, чтобы длина вектора $\mathbf{n} = \lambda \mathbf{N}$ была равна единице. Тогда

$$(\lambda \mathbf{N}, \lambda \mathbf{N}) = 1, \quad \lambda^2 (\mathbf{N}, \mathbf{N}) = 1, \quad \lambda = \pm \frac{1}{\sqrt{(\mathbf{N}, \mathbf{N})}}.$$

Знак числа λ берем таким, чтобы произведение λD было отрицательным. Обозначим $\lambda D = -p$, где $p > 0$. При этом уравнение (6) запишется в виде:

$$(\mathbf{n}, \mathbf{r}) - p = 0. \quad (7)$$

¹⁾ Предполагается, что коэффициенты A , B и C одновременно не обращаются в нуль.

Как было показано в § 17, это уравнение есть уравнение плоскости, имеющей нормальный вектор $\mathbf{n}$ и находящейся от начала на расстоянии p .

Таким образом, данное уравнение первой степени относительно x, y, z :

$$Ax + By + Cz + D = 0,$$

есть уравнение плоскости.

Для приведения этого уравнения к нормальному виду достаточно умножить обе части уравнения на множитель λ (нормирующий множитель):

$$\lambda = \frac{1}{\pm \sqrt{N, N}} = \pm \frac{1}{\sqrt{A^2 + B^2 + C^2}},$$

выбрав перед корнем знак, противоположный знаку числа D .

Уравнение $(N, \mathbf{r}) + D = 0$ называется *общим уравнением плоскости в векторной форме*.

Так как вектор N коллинеарен вектору $\mathbf{n}$, то вектор $N\{A, B, C\}$ расположен перпендикулярно к плоскости $(N, \mathbf{r}) + D = 0$, т. е. к плоскости

$$Ax + By + Cz + D = 0.$$

Уравнение $Ax + By + Cz + D = 0$ называется *общим уравнением плоскости*. Приведем его к нормальному виду:

$$\frac{A}{\pm \sqrt{A^2 + B^2 + C^2}} x + \frac{B}{\pm \sqrt{A^2 + B^2 + C^2}} y + \frac{C}{\pm \sqrt{A^2 + B^2 + C^2}} z + \frac{D}{\pm \sqrt{A^2 + B^2 + C^2}} = 0. \quad (8)$$

Полученное уравнение должно иметь вид:

$$x \cos \alpha + y \cos \beta + z \cos \gamma - p = 0.$$

Отсюда можно найти выражения для направляющих косинусов вектора N и расстояние p плоскости от начала координат:

$$\cos \alpha = \frac{A}{\pm \sqrt{A^2 + B^2 + C^2}}, \quad (9)$$

$$\cos \beta = \frac{B}{\pm \sqrt{A^2 + B^2 + C^2}}, \quad (10)$$

$$\cos \gamma = \frac{C}{\pm \sqrt{A^2 + B^2 + C^2}}, \quad (11)$$

$$p = \frac{-D}{\pm \sqrt{A^2 + B^2 + C^2}} \quad (12)$$

(знак перед корнем противоположен знаку D).

§ 16. Исследование общего уравнения плоскости

Рассмотрим некоторые частные случаи уравнения плоскости:

$$Ax + By + Cz + D = 0.$$

1. $D = 0$. Уравнение примет вид $Ax + By + Cz = 0$. Из равенства (12) § 15 следует, что плоскость проходит через начало координат (что видно и непосредственно).

2. $C = 0$. Уравнение принимает вид $Ax + By + D = 0$. На основании равенства (11) § 15 можно утверждать, что угол между нормалью к плоскости и осью OZ равен $\frac{\pi}{2}$, следовательно, плоскость $Ax + By + D = 0$ параллельна оси OZ . Аналогично заключаем, что плоскость $Ax + Cz + D = 0$ параллельна оси OY и плоскость $By + Cz + D = 0$ параллельна оси OX .

3. $C=0$ и $D=0$. Уравнение принимает вид $Ax + By = 0$. На основании равенств (11), (12) § 15 можно утверждать, что эта плоскость проходит через начало координат и параллельна оси OZ , т. е. $Ax + By = 0$ есть уравнение плоскости, проходящей через ось OZ . Аналогично получаем, что $Ax + Cz = 0$ — уравнение плоскости, проходящей через ось OY , а $By + Cz = 0$ — уравнение плоскости, проходящей через ось OX .

4. $B=0$, $C=0$. Уравнение принимает вид $Ax + D = 0$ или $x = -\frac{D}{A}$. Это уравнение плоскости, одновременно параллельной оси OZ и оси OY , т. е. параллельной координатной плоскости YOZ , аналогично $y = -\frac{D}{B}$ есть уравнение плоскости, параллельной координатной плоскости XOZ , а уравнение $z = -\frac{D}{C}$ есть уравнение плоскости, параллельной плоскости XOY .

5. $B=0$, $C=0$ и $D=0$, $A \neq 0$. Уравнение принимает вид $Ax = 0$ или $x = 0$. Это уравнение координатной плоскости YOZ . Аналогично $y = 0$ — уравнение координатной плоскости ZOX , уравнение $z = 0$ — уравнение координатной плоскости XOY .

§ 17. Связка плоскостей и ее приложения

Через всякую точку пространства можно провести бесчисленное множество плоскостей. Совокупность плоскостей, проходящих через фиксированную точку M_0 , называется *связкой плоскостей* с центром M_0 .

Возьмем уравнение плоскости в общем виде:

$$(N, r) + D = 0, \quad (1)$$

где $N\{A, B, C\}$; $r\{x, y, z\}$, и фиксированную точку M_0 с радиусом-вектором $r_0\{x_0, y_0, z_0\}$.

Пусть плоскость (1) проходит через точку M_0 . Тогда, подставляя в уравнение плоскости вместо текущего радиуса-вектора r фиксированный радиус-вектор r_0 , получим тождество:

$$(N, r_0) + D = 0. \quad (2)$$

Вычитая из равенства (1) тождество (2), получим:

$$(N, (r - r_0)) = 0. \quad (3)$$

При всяком ненулевом векторе N это будет уравнение некоторой плоскости, проходящей через точку M_0 , и уравнение всякой плоскости, проходящей через точку M_0 , можно записать в таком виде.

Уравнение (3) называется *уравнением связки плоскостей*. Так как вектор $(r - r_0)$ имеет координаты $x - x_0$, $y - y_0$, $z - z_0$, то в координатной форме уравнение связки запишется в виде:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \quad (4)$$

Уравнение плоскости, проходящей через три точки. Составим уравнение плоскости, проходящей через точки M_1 , M_2 , M_3 , определенные соответственно радиусами-векторами:

$$r_1\{x_1, y_1, z_1\}; \quad r_2\{x_2, y_2, z_2\}; \quad r_3\{x_3, y_3, z_3\}.$$

Для этого воспользуемся уравнением связки плоскостей с центром в точке M_1 :

$$(N, (r - r_1)) = 0 \quad (5)$$

и будем из этой связки плоскостей выбирать такую плоскость, которая проходила бы через точки M_2 и M_3 . Это приведет к соотношениям:

$$(N, (r_2 - r_1)) = 0, \quad (6)$$

$$(N, (r_3 - r_1)) = 0. \quad (7)$$

Каждое из равенств (5), (6), (7) представляет собой условие ортогональности векторов: $\mathbf{r} - \mathbf{r}_1$; $\mathbf{r}_2 - \mathbf{r}_1$; $\mathbf{r}_3 - \mathbf{r}_1$ к вектору $\mathbf{N}$:

$$\begin{aligned} \mathbf{N} &\perp (\mathbf{r} - \mathbf{r}_1), \\ \mathbf{N} &\perp (\mathbf{r}_2 - \mathbf{r}_1), \\ \mathbf{N} &\perp (\mathbf{r}_3 - \mathbf{r}_1). \end{aligned} \quad (8)$$

Но если три вектора ортогональны одному и тому же вектору, то эти векторы компланарны. Условием компланарности трех векторов является равенство нулю их смешанного произведения. Следовательно,

$$(\mathbf{r} - \mathbf{r}_1)(\mathbf{r}_2 - \mathbf{r}_1)(\mathbf{r}_3 - \mathbf{r}_1) = 0. \quad (9)$$

Уравнение (9) является *уравнением плоскости, проходящей через три точки*. В координатной форме это уравнение будет иметь вид:

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0. \quad (10)$$

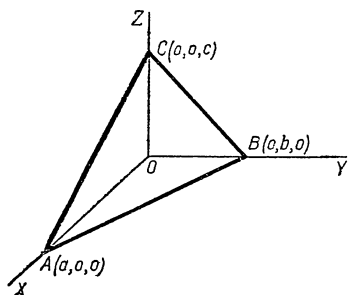


Рис. 87.

Пример. Составить уравнение плоскости, проходящей через три точки $M_1(1, 2, -1)$, $M_2(2, 2, 3)$, $M_3(1, 0, -2)$. Используя уравнение (10), получим:

$$\begin{vmatrix} x-1 & y-2 & z+1 \\ 1 & 0 & 4 \\ 0 & -2 & -1 \end{vmatrix} = 0, \text{ т. е. } 8x + y - 2z - 12 = 0.$$

Уравнение плоскости в отрезках на осях. Найдём уравнение плоскости, отсекающей на координатных осях отрезки, величины которых равны: a , b , c .

Очевидно, что точками пересечения плоскости с координатными осями будут точки $(a, 0, 0)$, $(0, b, 0)$, $(0, 0, c)$ (рис. 87).

Используя уравнение плоскости, проходящей через три точки, получим:

$$\begin{vmatrix} x-a & y & z \\ -a & b & 0 \\ -a & 0 & c \end{vmatrix} = 0, \\ bc(x-a) + acy + abz = 0, \\ bcx + acy + abz = abc, \\ \frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \quad (11)$$

Уравнение (11) называется *уравнением плоскости в отрезках на осях*.

§ 18. Расстояние от точки до плоскости

Пусть дана плоскость P своим уравнением в нормальном виде

$$(\mathbf{n}, \mathbf{r}) - p = 0,$$

где $\mathbf{n} \{ \cos \alpha, \cos \beta, \cos \gamma \}$, $\mathbf{r} \{ x, y, z \}$, и точка M_0 с радиусом-вектором $\mathbf{r}_0$. Требуется определить расстояние от точки M_0 до плоскости. Будем рассматривать случай, когда точка M_0 и начало координат O лежат по разные стороны от плоскости. Проведем вектор $\overline{OL}$, перпендикулярный плоскости, опустим перпендикуляр M_0K из точки M_0 на плоскость P и введем еще два вектора: $\overline{M_0K}$ и $\overline{OK}$ (рис. 88).

Вектор $\overrightarrow{OK}$ можно рассматривать как сумму двух векторов:

$$\overrightarrow{OK} = \overrightarrow{OM_0} + \overrightarrow{M_0K}. \quad (1)$$

Обозначив длину вектора $\overrightarrow{M_0K}$ через d и замечая, что вектор $\overrightarrow{M_0K}$ коллинеарен единичному вектору $\mathbf{n}$, но противоположен ему по направлению, можем написать:

$$\overrightarrow{M_0K} = -\mathbf{n}d, \quad (2)$$

тогда

$$\overrightarrow{OK} = \mathbf{r}_0 - \mathbf{n}d. \quad (3)$$

Так как $\overrightarrow{OK}$ есть радиус-вектор точки K , лежащей на плоскости, то можно в уравнение плоскости вместо текущего вектора $\mathbf{r}$ подставить вектор $\overrightarrow{OK}$:

$$(\mathbf{n}(\mathbf{r}_0 - \mathbf{n}d)) - p = 0$$

или

$$(\mathbf{n}, \mathbf{r}_0) - d - p = 0,$$

откуда

$$d = (\mathbf{n}, \mathbf{r}_0) - p. \quad (4)$$

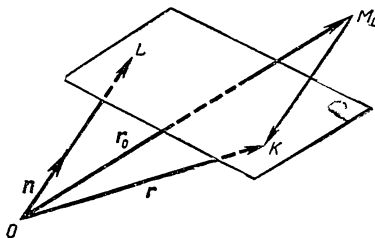


Рис. 88.

Равенство (4) позволяет определить расстояние от точки M_0 до плоскости P . Формула (4) была выведена в предположении, что точка M_0 и начало координат лежат по разные стороны от плоскости.

Если данная точка и начало координат лежат по одну сторону от плоскости (рис. 89), то, проводя аналогичные рассуждения, получим:

$$d = -(\mathbf{n}, \mathbf{r}_0) + p.$$

Таким образом, в общем случае:

$$d = |(\mathbf{n}, \mathbf{r}_0) - p|. \quad (5)$$

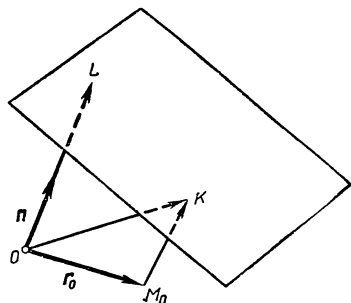


Рис. 89.

Этот результат можно изложить так: чтобы найти расстояние от точки до плоскости, надо уравнение плоскости привести к нормальному виду и взять абсолютную величину левой части уравнения, подставив в нее вместо радиуса-вектора произвольной точки плоскости радиус-вектор данной точки.

Если точка M_0 имеет координаты x_0, y_0, z_0 , а уравнение плоскости P дается в виде: $Ax + By + Cz + D = 0$, то нормируя уравнение, а затем подставляя вместо текущих координат координаты точки M_0 , получим формулу для определения расстояния от точки M_0 до плоскости P в координатной форме:

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}. \quad (6)$$

Пример. Найти расстояние от точки $M(5, 3, 2)$ до плоскости

$$2x - 3y + 6z + 4 = 0.$$

По формуле (6)

$$d = \frac{|2 \cdot 5 - 3(-3) + 6 \cdot 2 + 4|}{\sqrt{4 + 9 + 36}} = 5.$$

Таким образом, расстояние от точки до плоскости равно 5 единицам.

§ 19. Угол между двумя плоскостями

Пусть даны две плоскости своими уравнениями $(N_1, r) + D_1 = 0 \dots (P_1)$, $(N_2, r) + D_2 = 0 \dots (P_2)$, где векторы $N_1 \{A_1, B_1, C_1\}$ и $N_2 \{A_2, B_2, C_2\}$ соответственно перпендикулярны заданным плоскостям (рис. 90).

Углом φ между двумя плоскостями будем называть любой из двух смежных двугранных углов, образованных этими плоскостями¹⁾.

Обозначим угол между векторами N_1 и N_2 через θ .

Тогда $\cos \theta = \frac{(N_1, N_2)}{\|N_1\| \|N_2\|}$. Углы φ и θ , как следует из геометрических соображений, связаны равенствами $\varphi = \theta$ или $\varphi = \pi - \theta$, следовательно, $\cos \varphi = \cos \theta$ или $\cos \varphi = -\cos \theta$.

Отсюда $\cos \varphi = \pm \cos \theta$, т. е.

$$\cos \varphi = \pm \frac{(N_1, N_2)}{\|N_1\| \|N_2\|}. \quad (1)$$

Равенство (1) позволяет определить угол между плоскостями, заданными в векторной форме.

Если плоскости заданы уравнениями $A_1x + B_1y + C_1z + D_1 = 0$ и $A_2x + B_2y + C_2z + D_2 = 0$, то формула (1) запишется в виде:

$$\cos \varphi = \pm \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2} \sqrt{A_2^2 + B_2^2 + C_2^2}}. \quad (2)$$

Условия параллельности и перпендикулярности плоскостей. Для того чтобы плоскости P_1 и P_2 были параллельны, необходимо и достаточно, чтобы векторы $N_1 \{A_1, B_1, C_1\}$ и $N_2 \{A_2, B_2, C_2\}$ были

коллинеарны. В случае коллинеарности векторов их соответствующие координаты пропорциональны:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}. \quad (3)$$

Равенство (3) и является *условием параллельности* двух плоскостей.

Для того чтобы плоскости P_1 и P_2 были перпендикулярны, необходимо и достаточно, чтобы векторы N_1 и N_2 были ортогональны друг другу. В случае ортогональности векторов их координаты связаны соотношением:

$$A_1A_2 + B_1B_2 + C_1C_2 = 0. \quad (4)$$

Это равенство и является *условием перпендикулярности* двух плоскостей.

Пример 1. Составить уравнение плоскости, проходящей через точку $M(2, 3, 1)$ параллельно плоскости

$$5x - 4y + 7z - 43 = 0.$$

Составим уравнение связки плоскостей с центром в точке M :

$$A(x - 2) + B(y + 3) + C(z - 1) = 0.$$

На основании условия (3) параллельности плоскостей можно написать

$$\frac{A}{5} = \frac{B}{-4} = \frac{C}{7} = t$$

или

$$A = 5t; \quad B = -4t; \quad C = 7t.$$

¹⁾ Если плоскости параллельны, то угол между ними считаем равным нулю (или π).

²⁾ См. сноску на стр. 102.

Подставим в уравнение связки плоскостей значения A, B, C :

$$5t(x-2) - 4t(y+3) + 7t(z-1) = 0.$$

Разделим обе части уравнения на t :

$$5(x-2) - 4(y+3) + 7(z-1) = 0.$$

После алгебраического преобразования искомое уравнение плоскости будет иметь вид:

$$5x - 4y + 7z - 29 = 0.$$

Примечание. Так как обе части уравнения можно сократить на коэффициент пропорциональности, то в дальнейшем мы будем в уравнение связки вместо неизвестных коэффициентов подставлять сразу величины, им пропорциональные.

Пример 2. Составить уравнение плоскости, проходящей через точку $M(3, -4, 1)$ перпендикулярно двум плоскостям:

$$2x - 3y + 4z - 17 = 0 \dots (P_1); \quad 5x - y + 2z + 35 = 0 \dots (P_2).$$

Составим уравнение связки плоскостей с центром в точке $M(3, -4, 1)$:

$$A(x-3) + B(y+4) + C(z-1) = 0.$$

На основании условий перпендикулярности искомой плоскости к плоскостям P_1 и P_2 можно записать следующие равенства:

$$2A - 3B + 4C = 0; \quad 5A - B + 2C = 0.$$

Из полученной однородной системы двух линейных уравнений с тремя неизвестными A, B, C можно определить их отношение

$$A:B:C = \begin{vmatrix} -3 & 4 \\ -1 & 2 \end{vmatrix} : \begin{vmatrix} 4 & 2 \\ 2 & 5 \end{vmatrix} : \begin{vmatrix} 2 & -3 \\ 5 & -1 \end{vmatrix} = -2:16:13.$$

А тогда искомое уравнение примет вид:

$$\begin{aligned} 2(x-3) - 16(y+4) - 13(z-1) &= 0, \\ 2x - 16y - 13z - 57 &= 0. \end{aligned}$$

Раздел 6

ПРЯМАЯ ЛИНИЯ В ПРОСТРАНСТВЕ

§ 20. Параметрические и канонические уравнения прямой линии в пространстве

Параметрические уравнения прямой линии. Возьмем произвольную прямую и вектор $s\{m, n, p\}$, параллельный данной прямой (рис. 91). Вектор s называется *направляющим вектором* прямой. Затем на прямой возьмем некоторую фиксированную точку M_0 и произвольную точку M . Обозначим через r_0 и r радиусы-векторы точек M_0 и M . Очевидно, что три вектора r_0 , r и $\overrightarrow{M_0M}$ связаны соотношением

$$r - r_0 = \overrightarrow{M_0M}. \quad (1)$$

Вектор $\overrightarrow{M_0M}$, коллинеарный вектору s , может быть представлен как st , т. е. $\overrightarrow{M_0M} = st$, где t является параметром, принимающим различные значения в зависимости от положения точки M на прямой линии. Уравнение (1) можно записать в виде:

$$r - r_0 = st$$

или

$$r = r_0 + st. \quad (2)$$

Так как этому уравнению удовлетворяют координаты любой точки M прямой и не удовлетворяют координаты точки, не лежащей на этой

прямой, то уравнение (2) является уравнением прямой линии в пространстве. Полученное уравнение называется *параметрическим уравнением прямой* в пространстве в векторной форме.

Если точка M имеет координаты x , y и z , а точка M_0 — координаты x_0 , y_0 , z_0 ,

$$\text{то } \mathbf{r} = \mathbf{r}\{x, y, z\}, \quad \mathbf{r}_0 = \mathbf{r}_0\{x_0, y_0, z_0\}.$$

Уравнение (2) можно при этом записать в виде:

$$xi + yj + zk = (x_0 + mt) \mathbf{i} + (y_0 + nt) \mathbf{j} + (z_0 + pt) \mathbf{k}. \quad (3)$$

Так как у равных векторов соответствующие координаты равны, то

$$\begin{cases} x = x_0 + mt, \\ y = y_0 + nt, \\ z = z_0 + pt. \end{cases} \quad (4)$$

Равенства (4) являются *параметрическими уравнениями прямой* в пространстве в *координатной форме*.

Канонические уравнения прямой линии. Исключая из равенств (4) параметр t , получим:

$$\frac{x - x_0}{m} = \frac{y - y_0}{n} = \frac{z - z_0}{p}^1). \quad (5)$$

Этим равенствам будут удовлетворять координаты каждой точки, лежащей на прямой, и только они, т. е. равенства (5) являются уравне-

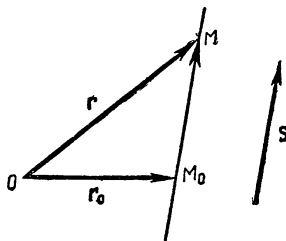


Рис. 91.

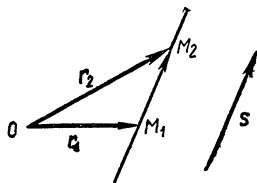


Рис. 92.

ниями прямой линии в пространстве. Их называют *каноническими уравнениями прямой*, или *уравнениями прямой в форме пропорций*.

Направляющими косинусами прямой называются направляющие косинусы вектора $\mathbf{s}$. Используя формулы (3) § 5, получим:

$$\begin{aligned} \cos \alpha &= \frac{m}{\sqrt{m^2 + n^2 + p^2}}, \\ \cos \beta &= \frac{n}{\sqrt{m^2 + n^2 + p^2}}, \\ \cos \gamma &= \frac{p}{\sqrt{m^2 + n^2 + p^2}}. \end{aligned} \quad (6)$$

Из равенств (6) определяем:

$$m : n : p = \cos \alpha : \cos \beta : \cos \gamma. \quad (7)$$

Следовательно, числа m , n и p пропорциональны направляющим косинусам прямой. Их принято называть *угловыми коэффициентами прямой* в пространстве. В канонических уравнениях прямой (5) x , y и z являются координатами произвольной точки, а x_0 , y_0 и z_0 координатами фиксированной точки M_0 , лежащей на прямой.

¹⁾ См. сноску на стр. 102.

Заметим, что числа m , n и p одновременно обращаются в нуль не могут, так как вектор s не нулевой. Однако один или два угловых коэффициента одновременно могут обращаться в нуль. Если, например, $m=0$, то $\cos \alpha = 0$, т. е. данная прямая перпендикулярна к оси абсцисс, и т. д.

Уравнения прямой в пространстве, проходящей через две точки. Пусть дана некоторая прямая, проходящая через точку M_1 , радиус-вектор которой $r_1 \{x_1, y_1, z_1\}$, и через точку M_2 , радиус-вектор которой $r_2 \{x_2, y_2, z_2\}$. Требуется составить уравнения этой прямой.

Составим уравнения прямой, проходящей через точку $M_1(x_1, y_1, z_1)$ (рис. 92), в каноническом виде:

$$\frac{x-x_1}{m} = \frac{y-y_1}{n} = \frac{z-z_1}{p}, \quad (8)$$

где $s \{m, n, p\}$ — направляющий вектор прямой.

Рассматривая векторы r_1 , r_2 и $\overline{M_1M_2} = r_2 - r_1$. Координатами вектора $r_2 - r_1$ будут $x_2 - x_1$, $y_2 - y_1$, $z_2 - z_1$.

Так как вектор $r_2 - r_1$ коллинеарен вектору s , то их координаты пропорциональны, т. е.

$$\frac{x_2 - x_1}{m} = \frac{y_2 - y_1}{n} = \frac{z_2 - z_1}{p}. \quad (9)$$

Подставив в уравнение (8) вместо m , n и p величины, им пропорциональные, получим искомые *уравнения прямой, проходящей через две точки*:

$$\frac{x-x_1}{x_2-x_1} = \frac{y-y_1}{y_2-y_1} = \frac{z-z_1}{z_2-z_1}^1). \quad (10)$$

§ 21. Общие уравнения прямой в пространстве

Пусть заданы уравнения двух пересекающихся плоскостей:

$$(N_1, r) + D_1 = 0 \text{ и } (N_2, r) + D_2 = 0$$

где

$$N_1 \{A_1, B_1, C_1\}; \quad N_2 \{A_2, B_2, C_2\}; \quad r \{x, y, z\}.$$

Тогда совокупность этих уравнений:

$$\begin{cases} (N_1, r) + D_1 = 0, \\ (N_2, r) + D_2 = 0 \end{cases} \quad (1)$$

можно рассматривать как уравнения прямой — линии пересечения этих плоскостей. Уравнения (1) называются *общими уравнениями прямой в пространстве в векторной форме*.

Выразив уравнения (1) в координатной форме, получим:

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \quad (2)$$

Уравнения (2) называются *общими уравнениями прямой в координатной форме*.

Приведение общих уравнений прямой к каноническому виду. Пусть даются уравнения прямой в общем виде:

$$\begin{cases} (N_1, r) + D_1 = 0, \\ (N_2, r) + D_2 = 0. \end{cases} \quad (2)$$

Уравнение $(N_1, r) + D_1 = 0$ есть уравнение плоскости P_1 , перпендикулярной вектору $N_1 \{A_1, B_1, C_1\}$; $(N_2, r) + D_2 = 0$ — уравнение плоскости P_2 , перпендикулярной вектору $N_2 \{A_2, B_2, C_2\}$ (рис. 93). Уравнение ли-

¹⁾ См. сноску к стр. 102.

нии их пересечения можно записать в векторной форме: $r = r_0 + st$, где $s \{m, n, p\}$ — направляющий вектор этой линии.

Найдем векторное произведение $[N_1, N_2]$:

$$[N_1, N_2] = \begin{vmatrix} i & j & k \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix} = i \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} + j \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix} + k \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}. \quad (3)$$

Из определения векторного произведения следует, что вектор $[N_1, N_2]$ коллинеарен вектору s . Следовательно, координаты этих двух векторов пропорциональны:

$$m : n : p = \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} : \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix} : \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}. \quad (4)$$

Таким образом, в канонические уравнения прямой

$$\frac{x-x_0}{m} = \frac{y-y_0}{n} = \frac{z-z_0}{p}$$

вместо коэффициентов m, n, p можно подставить числа, им пропорциональные, и получить уравнения

$$\frac{x-x_0}{\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}} = \frac{y-y_0}{\begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}} = \frac{z-z_0}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}. \quad (5)$$

Итак, чтобы от общих уравнений прямой

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases}$$

перейти к каноническим уравнениям $\frac{x-x_0}{m} = \frac{y-y_0}{n} = \frac{z-z_0}{p}$ той же прямой, надо найти

какую-либо точку $M_0(x_0, y_0, z_0)$, лежащую на прямой, и вместо коэффициентов m, n и p подставить пропорциональные им числа (см. (4)).

Следует заметить, что при нахождении точки, лежащей на прямой, одну из координат можно задавать совершенно произвольно; целесообразно за точку M_0 брать точку пересечения прямой с одной из координатных плоскостей, так как тогда хотя бы одна из координат этой точки будет равна нулю.

Пример. Уравнения прямой

$$\begin{cases} 2x - 3y + 5z + 7 = 0, \\ x + 3y - 4z - 1 = 0 \end{cases}$$

привести к каноническому виду.

За точку $M_0(x_0, y_0, z_0)$ возьмем точку пересечения этой прямой с плоскостью XOY , тогда $z_0 = 0$. Для определения x_0, y_0 будем иметь следующую систему:

$$\begin{cases} 2x_0 - 3y_0 + 7 = 0, \\ x_0 + 3y_0 - 1 = 0. \end{cases}$$

Решая эту систему, найдем, что $x_0 = -2$; $y_0 = 1$.

Используя равенства (4) для нахождения отношения коэффициентов m, n и p , получим:

$$m : n : p = \begin{vmatrix} -3 & 5 \\ 3 & -4 \end{vmatrix} : \begin{vmatrix} 5 & 2 \\ -4 & 1 \end{vmatrix} : \begin{vmatrix} 2 & -3 \\ 1 & 3 \end{vmatrix} = -3 : 13 : 9.$$

Следовательно, искомые уравнения имеют вид:

$$\frac{x+2}{-3} = \frac{y-1}{13} = \frac{z}{9}.$$

§ 22. Угол между двумя прямыми

Углом между двумя прямыми в пространстве называют любой из углов, образованных двумя прямыми, проведенными из одной точки параллельно данным прямым¹⁾.

Пусть даны уравнения двух прямых:

$$(l_1) \dots r = r_1 + s_1 t,$$

где $r \{x, y, z\}$, $r_1 \{x_1, y_1, z_1\}$, $s_1 \{m_1, n_1, p_1\}$, и

$$(l_2) \dots r = r_2 + s_2 t,$$

где $r_2 \{x_2, y_2, z_2\}$, $s_2 \{m_2, n_2, p_2\}$.

Обозначим угол между прямыми l_1 и l_2 (рис. 94) через φ , а угол между их направляющими векторами s_1 и s_2 — через θ . При этом

$$\cos \theta = \frac{(s_1, s_2)}{|s_1| |s_2|}. \quad (1)$$

Так как $\varphi = \theta$ или $\varphi = \pi - \theta$, то $\cos \varphi = \pm \cos \theta$. Следовательно,

$$\cos \varphi = \pm \frac{(s_1, s_2)}{|s_1| |s_2|}. \quad (2)$$

Если уравнения двух прямых даны в канонической форме:

$$\frac{x-x_1}{m_1} = \frac{y-y_1}{n_1} = \frac{z-z_1}{p_1} \dots (l_1);$$

$$\frac{x-x_2}{m_2} = \frac{y-y_2}{n_2} = \frac{z-z_2}{p_2} \dots (l_2),$$

то формулу (2) можно записать в координатной форме:

$$\cos \varphi = \pm \frac{m_1 m_2 + n_1 n_2 + p_1 p_2}{\sqrt{m_1^2 + n_1^2 + p_1^2} \sqrt{m_2^2 + n_2^2 + p_2^2}}. \quad (3)$$

Формулы (2) и (3) являются формулами для определения угла между двумя прямыми в пространстве.

Условия параллельности и перпендикулярности двух прямых в пространстве. Для того чтобы две прямые

$$r = r_1 + s_1 t \dots (l_1)$$

и

$$r = r_2 + s_2 t \dots (l_2),$$

где $s_1 \{m_1, n_1, p_1\}$ и $s_2 \{m_2, n_2, p_2\}$ были параллельны, необходимо и достаточно, чтобы их направляющие векторы s_1 и s_2 были коллинеарны, т. е. соответствующие координаты векторов s_1 и s_2 были пропорциональны:

$$\frac{m_1}{m_2} = \frac{n_1}{n_2} = \frac{p_1}{p_2}. \quad (4)$$

Условие (4) является условием параллельности двух прямых l_1 и l_2 в пространстве.

Для того чтобы прямые l_1 и l_2 были перпендикулярны между собой, необходимо и достаточно, чтобы параллельные им векторы s_1 и s_2 были ортогональными.

Условие ортогональности двух векторов s_1 и s_2 :

$$m_1 m_2 + n_1 n_2 + p_1 p_2 = 0 \quad (5)$$

является условием перпендикулярности двух прямых l_1 и l_2 в пространстве.

¹⁾ При этом, если прямые параллельны, угол между ними считается равным нулю (или π).

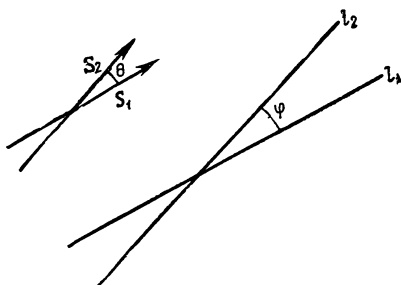


Рис. 94.

Пример 1. Найти уравнения прямой, проходящей через точку $M(9, -13, 15)$ перпендикулярно двум прямым l_1 и l_2 :

$$\frac{x-3}{2} = \frac{y+4}{-3} = \frac{z}{5} \dots (l_1); \quad \frac{x+2}{4} = \frac{y-7}{1} = \frac{z}{-2} \dots (l_2).$$

Составим уравнения любой прямой, проходящей через точку M :

$$\frac{x-9}{m} = \frac{y+13}{n} = \frac{z-15}{p}. \quad (6)$$

Используя условие перпендикулярности искомой прямой сначала к прямой l_1 , а затем к прямой l_2 , получим

$$\begin{aligned} 2m - 3n + 5p &= 0, \\ 4m + n - 2p &= 0. \end{aligned}$$

Из этой однородной системы линейных уравнений с неизвестными m, n и p найдем отношения неизвестных:

$$m:n:p = \begin{vmatrix} -3 & 5 \\ 1 & -2 \end{vmatrix} : \begin{vmatrix} 5 & 2 \\ -2 & 4 \end{vmatrix} : \begin{vmatrix} 2 & -3 \\ 4 & 1 \end{vmatrix} = 1:24:14.$$

Подставляя в уравнения прямой (6) вместо m, n и p пропорциональные им величины, получим искомые уравнения:

$$\frac{x-9}{1} = \frac{y+13}{24} = \frac{z-15}{14}.$$

Пример 2. Составить уравнения прямой, проходящей через точку $(8, -5, 0)$ параллельно прямой

$$l \dots \begin{cases} 2x - 3y + 5z - 17 = 0, \\ x + 4y - 2z + 8 = 0. \end{cases}$$

Составим уравнения любой прямой, проходящей через точку $M(8, -5, 0)$:

$$\frac{x-8}{m_1} = \frac{y+5}{n_1} = \frac{z}{p_1}. \quad (7)$$

Обозначим угловые коэффициенты прямой l через m, n, p и найдем их отношения, используя равенства (4) § 21:

$$m:n:p = \begin{vmatrix} -3 & 5 \\ 4 & -2 \end{vmatrix} : \begin{vmatrix} 5 & 2 \\ -2 & 1 \end{vmatrix} : \begin{vmatrix} 2 & -3 \\ 1 & 4 \end{vmatrix} = -14:9:11.$$

Так как у параллельных прямых угловые коэффициенты пропорциональны, то в уравнения (7) вместо m_1, n_1, p_1 можно подставить величины, им пропорциональные. Получим уравнения

$$\frac{x-8}{-14} = \frac{y+5}{9} = \frac{z}{11},$$

которые и будут уравнениями искомой прямой.

Раздел 7 ПРЯМАЯ И ПЛОСКОСТЬ

§ 23. Угол между прямой и плоскостью

Углом между прямой и плоскостью называют любой из двух смежных углов, образованных прямой и ее проекцией на плоскость. Пусть дано уравнение плоскости P :

$$(N, r) + D = 0,$$

где $r\{x, y, z\}$ — текущий вектор, а вектор $N\{A, B, C\}$ перпендикулярен к плоскости P , и уравнение прямой l :

$$r = r_0 + st,$$

где вектор $s\{m, n, p\}$ параллелен прямой l (рис. 95).

Обозначим угол между векторами N и s через θ , а угол между плоскостью P и прямой l — через φ . Найдем косинус угла θ между векторами N и s :

$$\cos \theta = \frac{(N, s)}{|N| |s|}.$$

При этом $\sin \varphi = \pm \cos \theta$. Следовательно,

$$\sin \varphi = \pm \frac{(N, s)}{|N| |s|} \quad (1)$$

или, в координатной форме,

$$\sin \varphi = \pm \frac{Am + Bn + Cp}{\sqrt{A^2 + B^2 + C^2} \sqrt{m^2 + n^2 + p^2}}.$$

Условия параллельности и перпендикулярности прямой и плоскости. Для того чтобы плоскость $(P) \dots Nr + D = 0$ была параллельна прямой $(l) \dots r = r_0 + st$, необходимо и достаточно, чтобы векторы $N \{A, B, C\}$ и $s \{m, n, p\}$ были ортогональны между собой.

Условие ортогональности двух векторов N и s может быть записано так:

$$(N, s) = 0$$

или так:

$$Am + Bn + Cp = 0. \quad (2)$$

Для того чтобы прямая l была перпендикулярна плоскости P , необходимо и достаточно, чтобы вектор s был коллинеарен вектору N .

Условие коллинеарности двух векторов N и s может быть записано следующим образом:

$$[N, s] = 0$$

или

$$\frac{A}{m} = \frac{B}{n} = \frac{C}{p}. \quad (3)$$

Если уравнение плоскости дается в общем виде:

$$Ax + By + Cz + D = 0,$$

а уравнения прямой — в каноническом виде:

$$\frac{x - x_0}{m} = \frac{y - y_0}{n} = \frac{z - z_0}{p},$$

то условием параллельности прямой и плоскости является равенство (2), а условиями перпендикулярности — равенства (3).

Пример. Составить уравнение плоскости P , проходящей через точку $M(2, -1, 3)$ параллельно двум прямым:

$$l_1 \dots \frac{x-2}{3} = \frac{y+17}{4} = \frac{z-8}{-5}$$

и

$$l_2 \dots \frac{x+8}{2} = \frac{y-14}{-3} = \frac{z}{1}.$$

Напишем уравнение связки плоскостей с центром в точке M :

$$A(x-2) + B(y+1) + C(z-3) = 0. \quad (4)$$

Используем условие параллельности (2) плоскости P к прямой l_1 , а затем к прямой l_2 :

$$3A + 4B - 5C = 0,$$

$$2A - 3B + C = 0.$$

Из этой системы однородных уравнений определим отношения коэффициентов A, B, C и затем в уравнение (4) вместо коэффициентов A, B, C подставим пропорциональные им величины:

$$A:B:C = \left| \begin{array}{cc|cc} 4 & -5 & -5 & 3 \\ -3 & 1 & 1 & 2 \end{array} \right| : \left| \begin{array}{cc} 3 & 4 \\ 2 & -3 \end{array} \right| = -11:-13:-17,$$

$$11(x-2) + 13(y+1) + 17(z-3) = 0,$$

$$11x + 13y + 17z - 60 = 0.$$

§ 24. Точка пересечения прямой с плоскостью

Даны уравнение $(N, r) + D = 0$ плоскости P , где $N\{A, B, C\}$, $r\{x, y, z\}$, и уравнение $r = r_0 + st$ прямой l , где $s\{m, n, p\}$ и $r_0\{x_0, y_0, z_0\}$.

Пусть M_1 — точка пересечения прямой l с плоскостью P . Обозначим через t_1 значение параметра t для точки M_1 .

Так как при этом значении параметра вектор $r = r_0 + st_1$ является радиусом-вектором точки, лежащей на плоскости, то, подставив его в уравнение плоскости $(Nr) + D = 0$, мы получим: $(N, (r_0 + st_1)) + D = 0$, т. е.

$$t_1(N, s) + (N, r_0) + D = 0, \quad (1)$$

откуда, если $(N, s) \neq 0$, определим

$$t_1 = -\frac{(N, r_0) + D}{(N, s)}. \quad (2)$$

Подставляя значение t_1 в уравнение прямой l , найдем радиус-вектор r_1 точки пересечения прямой l с плоскостью P :

$$r_1 = r_0 + st_1. \quad (3)$$

Пусть теперь уравнения плоскости и прямой заданы в координатной форме:

$$(P) \dots Ax + By + Cz + D = 0,$$

$$(l) \dots \frac{x-x_0}{m} = \frac{y-y_0}{n} = \frac{z-z_0}{p}.$$

Тогда, используя равенство (2), получим:

$$t_1 = -\frac{Ax_0 + By_0 + Cz_0 + D}{Am + Bn + Cp}. \quad (4)$$

Затем, чтобы найти координаты точки $M_1(x_1, y_1, z_1)$ пересечения прямой с плоскостью, подставим найденное значение t_1 в параметрические уравнения прямой:

$$\begin{cases} x_1 = x_0 + mt_1, \\ y_1 = y_0 + nt_1, \\ z_1 = z_0 + pt_1. \end{cases} \quad (5)$$

Взаимное расположение прямой и плоскости в пространстве. Рассмотрим три случая взаимного расположения прямой и плоскости в пространстве.

I. $(N, s) \neq 0$ (или, в координатной форме, $Am + Bn + Cp \neq 0$). В этом случае параметр t_1 имеет единственное значение. Следовательно, прямая и плоскость пересекаются в единственной точке.

II. $(N, s)=0$; $(N, r_0)+D \neq 0$ (или, в координатной форме, $Am+Bn+Cp=0$; $Ax_0+By_0+Cz_0+D \neq 0$). В этом случае не существует значений t_1 , удовлетворяющих уравнению (1), т. е. прямая и плоскость не пересекаются. Условие $(N, s)=0$ ($Am+Bn+Cp=0$) является условием параллельности прямой и плоскости.

III. $(N, s)=0$; $(N, r_0)+D=0$ (или, в координатной форме, $Am+Bn+Cp=0$; $Ax_0+By_0+Cz_0+D=0$). В этом случае уравнению (1) удовлетворяет любое значение t_1 , т. е. прямая и плоскость имеют бесчисленное множество точек пересечения или, что то же, прямая лежит на плоскости.

Таким образом, условия

$$\begin{cases} (N, s)=0, \\ (N, r_0)+D=0. \end{cases} \quad \begin{pmatrix} Am+Bn+Cp=0 \\ Ax_0+By_0+Cz_0+D=0 \end{pmatrix} \quad (6)$$

можно назвать *условиями принадлежности прямой плоскости*.

Пример 1. Найти проекцию точки $M(9, -13, -18)$ на плоскость P

$$9x - 17y - 15z + 23 = 0. \quad (7)$$

Для решения задачи составим уравнения прямой, проходящей через точку M перпендикулярно плоскости P .

Уравнения всякой прямой, проходящей через точку M , будут иметь вид:

$$\frac{x-9}{m} = \frac{y+13}{n} = \frac{z+18}{p}.$$

На основании условий перпендикулярности прямой и плоскости $\left(\frac{A}{m} = \frac{B}{n} = \frac{C}{p}\right)$ можно в уравнения прямой вместо коэффициентов m, n, p подставить пропорциональные им величины A, B, C :

$$\frac{x-9}{9} = \frac{y+13}{-17} = \frac{z+18}{-15}. \quad (8)$$

Для нахождения проекции N точки M на плоскость надо найти точку пересечения плоскости (7) и прямой (8). По формуле (4) определяем

$$t_1 = -\frac{81 + 221 + 270 + 23}{81 + 289 + 225} = -1.$$

Отсюда координаты точки $N(x_N, y_N, z_N)$ определяются по формулам (5):

$$\begin{aligned} x_N &= 9 - 9 = 0, \\ y_N &= -13 + 17 = 4, \\ z_N &= -18 + 15 = -3, \end{aligned}$$

т. е. $N(0, 4, -3)$.

Пример 2. Найти расстояние от точки $M(-2, 3, 4)$ до прямой

$$(l) \dots \frac{x-3}{1} = \frac{y+1}{2} = \frac{z-1}{1}.$$

План решения примера следующий:

1. Составим уравнение плоскости P , проходящей через точку M перпендикулярно прямой l .

2. Найдем точку K пересечения прямой l с плоскостью P .

3. Определим расстояние от точки M до прямой l , как расстояние между точками M и K .

Решение. Составим уравнение связки плоскостей с центром в точке M :

$$A(x+2) + B(y-3) + C(z-4) = 0.$$

Вследствие перпендикулярности плоскости P и прямой l можно написать уравнение плоскости P :

$$(x+2) + 2(y-3) + (z-4) = 0, \quad x + 2y + z - 8 = 0.$$

Для точки K

$$t_1 = -\frac{3-2+1-8}{1+4+1} = 1,$$

откуда

$$\begin{aligned} x_k &= 3 + 1 = 4, \\ y_k &= -1 + 2 = 1, \\ z_k &= 1 + 1 = 2. \end{aligned}$$

$$MK = \sqrt{(4+2)^2 + (1-3)^2 + (2-4)^2} = 2\sqrt{11}.$$

Заметим, что точка K является проекцией точки M на прямую l .

§ 25. Пучок плоскостей

Через всякую прямую в пространстве можно провести бесчисленное множество плоскостей. Совокупность всех плоскостей, проходящих через одну и ту же прямую, называется *пучком плоскостей*.

Пусть дано уравнение прямой как линии пересечения двух плоскостей

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \quad (1)$$

Составим уравнение:

$$A_1x + B_1y + C_1z + D_1 + \lambda(A_2x + B_2y + C_2z + D_2) = 0, \quad (2)$$

где λ — произвольное число.

Прежде всего, очевидно, что при любом значении λ уравнение (2) является уравнением плоскости, так как оно является уравнением первой степени.

Далее, каждая из плоскостей, определяемых при различных значениях λ уравнением (2), проходит через прямую, заданную уравнениями (1).

Действительно, возьмем любую точку $M_1(x_1, y_1, z_1)$, лежащую на прямой, заданной уравнениями (1), и любую плоскость (2). Подставляя в уравнения (1) вместо текущих координат координаты точки M_1 , получим:

$$\begin{cases} A_1x_1 + B_1y_1 + C_1z_1 + D_1 = 0, \\ A_2x_1 + B_2y_1 + C_2z_1 + D_2 = 0, \end{cases} \quad (3)$$

и следовательно,

$$A_1x_1 + B_1y_1 + C_1z_1 + D_1 + \lambda(A_2x_1 + B_2y_1 + C_2z_1 + D_2) = 0. \quad (4)$$

Уравнение (2) называется *уравнением пучка плоскостей*, проходящих через прямую (1). Число λ иногда называют *параметром пучка*.

Следует сделать примечание, аналогичное примечанию к уравнению пучка прямых на плоскости: уравнение (2) дает любую плоскость пучка, за исключением плоскости

$$A_2x + B_2y + C_2z + D_2 = 0.$$

Пример 1. Найти проекцию прямой

$$l \dots \begin{cases} 2x - 3y + 4z - 1 = 0, \\ x + 5y - 2z + 3 = 0 \end{cases}$$

на плоскость $(P) \dots 3x - 4y + z - 8 = 0$.

Составим уравнение пучка плоскостей, проходящих через прямую l :

$$2x - 3y + 4z - 1 + \lambda(x + 5y - 2z + 3) = 0 \quad (5)$$

или

$$(2 + \lambda)x + (5\lambda - 3)y + (4 - 2\lambda)z + (-1 + 3\lambda) = 0. \quad (6)$$

Из этого пучка плоскостей выделим плоскость, проектирующую прямую l на плоскость P . Определим λ , используя условие перпендикулярности плоскостей:

$$3(2 + \lambda) - 4(5\lambda - 3) + (4 - 2\lambda) = 0,$$

откуда $\lambda = \frac{22}{19}$. Подставив значение λ в уравнение (5), найдем уравнение проектирующей плоскости Q :

$$2x - 3y + 4z - 1 + \frac{22}{19}(x + 5y - 2z + 3) = 0, \\ 60x + 53y + 32z + 47 = 0.$$

Уравнения искомой проекции l_1 можно записать как уравнения линии пересечения плоскостей P и Q :

$$l_1 \dots \begin{cases} 3x - 4y + x - 8 = 0, \\ 60x + 53y + 32z + 47 = 0. \end{cases}$$

Пример 2. Составить уравнение плоскости, проходящей через прямую

$$l \dots \begin{cases} 2x - 3y + 4z + 6 = 0, \\ x + 3y - 2z + 2 = 0, \end{cases}$$

и удаленной на расстояние в 2 единицы от начала координат.

Составим уравнение пучка плоскостей, проходящих через прямую l :

$$2x - 3y + 4z + 6 + \lambda(x + 3y - 2z + 2) = 0 \quad (7)$$

или

$$(2 + \lambda)x + (3\lambda - 3)y + (4 - 2\lambda)z + 6 + 2\lambda = 0.$$

Для определения λ используем формулу расстояния от точки до плоскости

$$d = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}.$$

По условию

$$d = 2; \quad x_1 = 0; \quad y_1 = 0; \quad z_1 = 0.$$

Следовательно,

$$2 = \frac{|6 + 2\lambda|}{\sqrt{(2 + \lambda)^2 + (3\lambda - 3)^2 + (4 - 2\lambda)^2}}, \\ 13\lambda^2 - 36\lambda + 20 = 0. \quad (8)$$

Для определения λ имеем квадратное уравнение. Если корни этого уравнения будут действительные и различные, то существуют две плоскости P_1 и P_2 , удовлетворяющие условию поставленной задачи; если корни будут действительные и равные, то имеется одна плоскость, удовлетворяющая условию; если корни окажутся мнимыми, то плоскости, удовлетворяющей условию, не существует.

Решая уравнение (8), находим, что $\lambda_1 = 2$; $\lambda_2 = \frac{10}{13}$. Подставляя поочередно в уравнение (7) значения λ_1 и λ_2 , найдем уравнения двух плоскостей P_1 и P_2 , удовлетворяющих условиям задачи:

$$4x + 3y + 10 = 0, \quad (P_1) \\ 36x - 9y + 32z + 98 = 0. \quad (P_2)$$

Раздел 8

ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА

В настоящем разделе будут в основном рассматриваться *поверхности второго порядка*, т. е. поверхности, уравнения которых в прямоугольной системе координат являются алгебраическими уравнениями второй степени.

При рассмотрении вопроса о поверхностях могут быть поставлены две задачи:

1. Зная геометрические свойства поверхности, составить ее уравнение.

2. Имея уравнение поверхности, определить ее геометрические свойства.

Рассмотрим основные задачи первого типа.

§ 26. Цилиндрические поверхности

Поверхность, образованная прямыми линиями, параллельными фиксированной прямой, называется *цилиндрической поверхностью*, или коротче — *цилиндром*. Упомянутые прямые, из которых образованы цилиндрические поверхности, называются *образующими* этой поверхности.

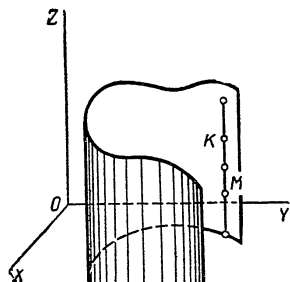


Рис. 96.

Цилиндрическая поверхность с образующими, параллельными данной прямой l , может быть получена так: через каждую точку некоторой линии C проводится прямая, параллельная прямой l , полученные прямые образуют цилиндрическую поверхность. Линия C называется *направляющей* этой поверхности. Очевидно, что всякая линия на цилиндрической поверхности, пересекающая все образующие этой поверхности, является ее направляющей.

Если уравнение поверхности в прямоугольных координатах не содержит z , т. е. имеет вид $F(x, y) = 0$, то эта поверхность цилиндрическая с образующими, параллельными оси OZ (рис. 96). В самом деле, если точка $M(x_0, y_0, z_0)$ принадлежит рассматриваемой поверхности, то для всякой точки $K(x, y, z)$, лежащей на прямой, параллельной OZ и проходящей через M , имеем: $x = x_0$; $y = y_0$; откуда следует, что координаты точки K удовлетворяют уравнению рассматриваемой поверхности. Это показывает, что рассматриваемая поверхность образована прямыми, параллельными оси OZ , и, следовательно, является цилиндрической с образующими, параллельными OZ .

Линия на плоскости XOY , имеющая уравнение $F(x, y) = 0$ (такое же, как уравнение рассматриваемой поверхности), является направляющей.

Аналогично уравнение вида $F(y, z) = 0$ есть уравнение цилиндрической поверхности с образующими, параллельными оси OX ; уравнение вида $F(x, z) = 0$ есть уравнение цилиндрической поверхности с образующими, параллельными оси OY .

Рассмотрим теперь цилиндрические поверхности второго порядка с образующими, параллельными оси OZ , заданные в прямоугольной системе координат XYZ уравнениями:

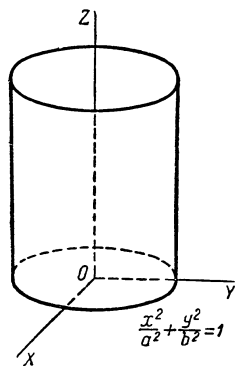


Рис. 97.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (1)$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad (2)$$

$$x^2 = 2py. \quad (3)$$

Направляющей цилиндрической поверхности (1) на плоскости XOY является эллипс, а сама поверхность $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ называется *эллиптичес-*

ким цилиндром (рис. 97). Направляющей цилиндрической поверхности (2) на плоскости XOY является гипербола, а сама поверхность $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ называется *гиперболическим цилиндром* (рис. 98).

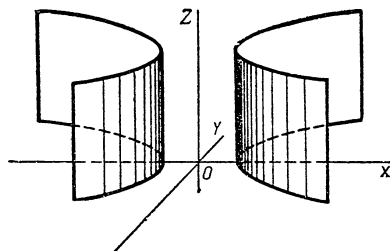


Рис. 98.

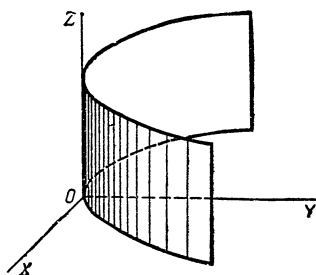


Рис. 99.

Направляющей цилиндрической поверхности (3) на плоскости XOY является парабола, а сама поверхность $x^2 = 2py$ называется *параболическим цилиндром* (рис. 99).

§ 27. Конические поверхности

Поверхность, образованная прямыми линиями, проходящими через фиксированную точку S , называется *конической поверхностью*, или короче — *конусом*. Упомянутые прямые, из которых образована коническая поверхность, называются *образующими* этой поверхности, а точка S — ее *вершиной*.

Коническая поверхность с вершиной S может быть получена так: через каждую точку некоторой линии C (не проходящей через S) и точку S проводится прямая; все эти прямые образуют коническую поверхность, причем линия C называется *направляющей* этой поверхности. Очевидно, что всякая линия на конической поверхности, пересекающая все образующие и не проходящая через вершину, является направляющей (рис. 100).

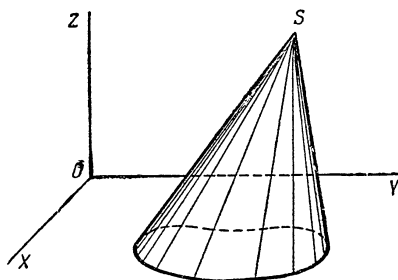


Рис. 100.

Если левая часть уравнения поверхности в прямоугольных координатах $F(x, y, z) = 0$ есть однородная функция относительно x, y, z какой-нибудь степени¹⁾, то поверхность будет конической с вершиной в начале O . В самом деле, если точка $M_0(x_0, y_0, z_0)$ принадлежит рассматриваемой поверхности, то всякая точка $M(x, y, z)$ лежащая на прямой, проходящей через точку M_0 и начало координат, также лежит на этой поверхности. Действительно, так как для этой точки имеем

¹⁾ Функция $F(x, y, z)$ называется *однородной n -й степени* относительно x, y, z , если для любого λ выполняется тождество $F(\lambda x, \lambda y, \lambda z) = \lambda^n F(x, y, z)$. Например, $F(x, y, z) = \frac{x^2}{2} - \frac{y^2}{3} + \frac{z^2}{4}$ является однородной функцией второй степени, так как

$$F(\lambda x, \lambda y, \lambda z) = \frac{\lambda^2 x^2}{2} - \frac{\lambda^2 y^2}{3} + \frac{\lambda^2 z^2}{4} = \lambda^2 \left(\frac{x^2}{2} - \frac{y^2}{3} + \frac{z^2}{4} \right) = \lambda^2 F(x, y, z).$$

$\Phi(x, y, z) = 5x^2y + xyz - 7yz^2$ является однородной функцией третьей степени и т. п. Подробнее о функциях многих переменных см. в главе VI.

$\frac{x}{x_0} = \frac{y}{y_0} = \frac{z}{z_0}$, т. е. $x = \lambda x_0$; $y = \lambda y_0$; $z = \lambda z_0$ (где λ — некоторое число) и, следовательно,

$$F(x, y, z) = F(\lambda x_0, \lambda y_0, \lambda z_0) = \lambda^n F(x_0, y_0, z_0) = \lambda^n \cdot 0 = 0$$

(где n — степень однородности функции F).

Это показывает, что рассматриваемая поверхность образована прямыми, проходящими через начало и, следовательно, является конической с вершиной в начале координат.

§ 28. Линейчатые поверхности

Цилиндрические и конические поверхности являются частными случаями так называемых *линейчатых поверхностей*, т. е. поверхностей, образованных прямыми линиями (эти прямые называются *прямолинейными образующими* линейчатой поверхности).

Среди поверхностей второго порядка не только цилиндры и конусы второго порядка являются линейчатыми поверхностями. Можно показать, что изучаемые ниже однополостные гиперболоиды и гиперболы параболоиды также являются линейчатыми поверхностями.

§ 29. Поверхности вращения

Поверхность, описываемая некоторой линией C , вращающейся вокруг неподвижной прямой d , называется *поверхностью вращения* с *осью вращения* d . Так как каждая точка линии C при вращении вокруг d описывает окружность, имеющую центр на прямой d и лежащую в плоскости, перпендикулярной к d , то можно сказать, что поверхность вращения с осью d есть поверхность, образованная окружностями, имеющими центры на оси d и лежащими в плоскостях, перпендикулярных d .

Если уравнение поверхности в прямоугольных координатах имеет вид $F(x^2 + y^2, z) = 0$, то эта поверхность будет поверхностью вращения с осью вращения OZ . В самом деле, если точка $M_0(x_0, y_0, z_0)$ лежит на рассматриваемой поверхности, то для всякой точки $M(x, y, z)$, принадлежащей окружности, проходящей через точку M_0 , имеющей центр на OZ и лежащей в плоскости, перпендикулярной OZ , имеют место равенства:

$$x^2 + y^2 = x_0^2 + y_0^2; \quad z = z_0,$$

откуда видно, что координаты точки M удовлетворяют уравнению рассматриваемой поверхности. Это показывает, что наша поверхность образована окружностями с центрами на оси OZ , лежащими в плоскостях, перпендикулярных OZ , и, следовательно, она является поверхностью вращения с осью OZ .

Аналогично уравнение вида $F(y^2 + z^2, x) = 0$ определяет поверхность вращения с осью вращения OX ; уравнение вида $F(x^2 + z^2, y) = 0$ определяет поверхность вращения с осью вращения OY .

Таким образом, поверхности второго порядка, заданные следующими уравнениями:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad (1)$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad (2)$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1, \quad (3)$$

$$\frac{x^2 + y^2}{p} = 2z, \quad (4)$$

являются поверхностями вращения с осью вращения OZ . Они соответственно называются: (1) — *эллипсоидом вращения*, (2) — *однополостным гипер-*

болоидом вращения, (3) — двухполостным гиперболоидом вращения, (4) — параболоидом вращения.

Эти поверхности являются частными случаями поверхностей второго порядка, форма которых будет изучена в дальнейшем в § 30.

§ 30. Метод сечений при исследовании формы поверхностей

Задачей предшествующих трех параграфов являлось составление уравнения поверхности по ее геометрическим свойствам. В настоящем параграфе будем решать другую задачу, а именно, по данному уравнению поверхности будем определять ее геометрический вид. Для этого применим так называемый «метод сечений». Метод сечений заключается в следующем: исследуя вид поверхности, заданной своим уравнением, будем искать линии пересечения данной поверхности с различными системами плоскостей, в частности с плоскостями, параллельными координатным плоскостям. Ряд таких сечений позволит определить вид поверхности.

Трехосный эллипсоид. Исследуем поверхность, заданную уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \quad (1)$$

Будем искать линии пересечения этой поверхности с плоскостями, параллельными плоскости $ХОУ$. Уравнения таких плоскостей:

$$z = h, \quad (2)$$

где h — любое число.

Решая совместно уравнения (1) и (2), получим уравнения

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{h^2}{c^2}; \quad z = h, \quad (3)$$

которые являются уравнениями линии пересечения поверхности с плоскостью $z = h$.

Исследуем уравнения (3) при различных значениях h .

1. Если $|h| > c$ ($c > 0$), то $1 - \frac{h^2}{c^2} < 0$, следовательно, $\frac{x^2}{a^2} + \frac{y^2}{b^2} < 0$, т. е. будем иметь уравнение мнимого места точек. Следовательно, точек пересечения поверхности с плоскостью $z = h$ не существует.

2. Если $|h| = c$, $h = \pm c$, то $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$. Линии пересечения (3) вырождаются в точки ($x = 0, y = 0, z = c$) и ($x = 0, y = 0, z = -c$). Плоскости $z = \pm c$ будут касаться данной поверхности.

3. Если $|h| < c$, то уравнения (3) можно переписать в виде:

$$\frac{x^2}{a^2 \left(1 - \frac{h^2}{c^2}\right)} + \frac{y^2}{b^2 \left(1 - \frac{h^2}{c^2}\right)} = 1; \quad z = h. \quad (4)$$

При указанных значениях h линия пересечения есть эллипс. Чем меньше будет $|h|$, тем с большими полуосями будут получаться эллипсы. Своего наибольшего значения полуоси достигнут тогда, когда $h = 0$. В этом случае уравнения (4) примут вид:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1; \quad h = 0,$$

т. е. в сечении поверхности координатной плоскостью $ХОУ$ получится эллипс с полуосями a и b . Аналогичные результаты получаются и при сечении данной поверхности плоскостями, параллельными координатным плоскостям $ХОZ$ и YOZ .

Полученные линии пересечения позволят нам изобразить поверхность, которая называется *трехосным эллипсоидом* (рис. 101). Рассмотренный

ранее эллипсоид вращения является частным случаем эллипсоида (1).

Однополостный гиперболоид. Исследуем поверхность, заданную уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \quad (5)$$

1. Ищем линии пересечения поверхности плоскостями $z=h$, параллельными плоскости XOY . Совместное решение уравнений (5) и (2) дает уравнения

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2}; \quad z=h \quad (6)$$

линии пересечения поверхности плоскостями $z=h$.

Запишем эти уравнения иначе:

$$\frac{x^2}{a^2 \left(1 + \frac{h^2}{c^2}\right)} + \frac{y^2}{b^2 \left(1 + \frac{h^2}{c^2}\right)} = 1; \quad z=h \quad (7)$$

Сечениями поверхности плоскостями, параллельными плоскости XOY , являются эллипсы, полуоси которых достигнут своего наименьшего значения тогда, когда $h=0$. По мере возрастания $|h|$, т. е. по мере удаления плоскостей от координатной плоскости XOY , полуоси эллипса будут увеличиваться.

2. Найдем линию пересечения поверхности (5) с плоскостью YOZ , урав-

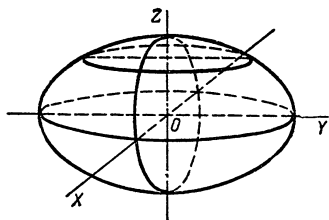


Рис. 101.

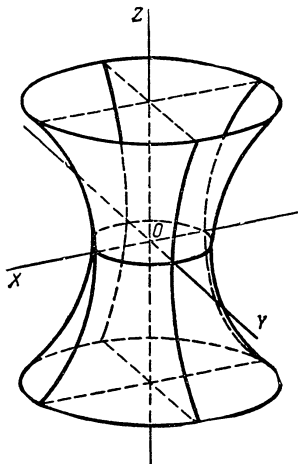


Рис. 102.

нение которой будет $x=0$. Совместное решение уравнений $x=0$ и (5) даст нам уравнения гиперболы:

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1; \quad x=0 \quad (8)$$

— линии пересечения данной поверхности с плоскостью YOZ .

3. Аналогично, находя линию пересечения поверхности (5) с плоскостью XOZ , получим уравнения гиперболы:

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1; \quad y=0.$$

Найденные линии пересечения позволяют изобразить поверхность, заданную уравнением $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$, эта поверхность называется *однополостным гиперболоидом* (рис. 102).

Рассмотренный ранее однополостный гиперболоид вращения является частным случаем однополостного гиперболоида (5).

Двухполостный гиперboloид. Исследуем поверхность, заданную уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1. \quad (9)$$

1. Ищем линии пересечения поверхности (9) с плоскостями $z=h$, параллельными плоскости XOY . Совместное решение уравнения (9) и уравнения плоскости $z=h$ дает:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{h^2}{c^2} = -1; \quad z=h$$

или

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{h^2}{c^2} - 1; \quad z=h. \quad (10)$$

При $|h| < c$ уравнения (10) являются уравнениями мнимого места точек, т. е. в этом случае плоскость не пересекает поверхность (9).

При $h = \pm c$ уравнения (10) удовлетворяются лишь координатами точек: $x=0, y=0, z=\pm c$; плоскости $z=\pm c$ касаются данной поверхности.

При $|h| > c$ уравнения (10) могут быть записаны так:

$$\frac{x^2}{a^2 \left(\frac{h^2}{c^2} - 1 \right)} + \frac{y^2}{b^2 \left(\frac{h^2}{c^2} - 1 \right)} = 1; \quad z=h. \quad (11)$$

Уравнения (11) являются уравнениями эллипса, полуоси которого по мере возрастания $|h|$ (т. е. удаления плоскости $z=|h|$ от координатной плоскости) увеличиваются.

2. Найдем теперь линию пересечения поверхности (9) с плоскостью YOZ , уравнение которой будет $x=0$.

Совместное решение этих уравнений показывает, что в сечении данной поверхности плоскостью YOZ лежит гипербола $\frac{y^2}{b^2} - \frac{z^2}{c^2} = -1; x=0$ (с действительной осью OZ).

3. Аналогично в сечении поверхности (9) плоскостью $y=0$ получается гипербола $\frac{x^2}{a^2} - \frac{z^2}{c^2} = -1; y=0$ (с действительной осью OZ).

Полученные линии пересечения позволяют изобразить поверхность

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1,$$

называемую *двухполостным гиперboloидом* (рис. 103).

Рассмотренный ранее двухполостный гиперboloид вращения является частным случаем двухполостного гиперboloида (9).

Эллиптический параболоид. Исследуем поверхность, заданную уравнением

$$\frac{x^2}{p} + \frac{y^2}{q} = 2z, \quad \text{где } p > 0; q > 0. \quad (12)$$

1. Найдем линии пересечения этой поверхности с плоскостями, параллельными плоскости XOY . Рассмотрим плоскость

$$z=h. \quad (13)$$

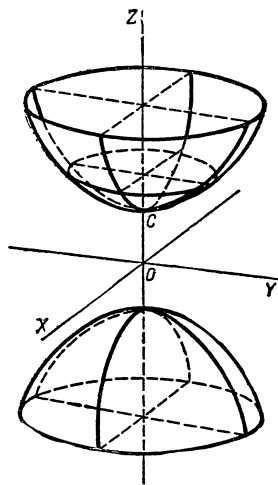


Рис. 103.

Совместное решение уравнений (12) и (13) дает уравнения

$$\frac{x^2}{p} + \frac{y^2}{q} = 2h; \quad z = h. \quad (14)$$

При $h < 0$ эти уравнения являются уравнениями мнимого места точек, т. е. при $h < 0$ плоскости $z = h$, параллельные плоскости XOY , данную поверхность не пересекают.

При $h = 0$ уравнения (14) удовлетворяются лишь координатами точки $(0, 0, 0)$; плоскость $z = 0$ касается данной поверхности.

При $h > 0$ уравнения (14) можно переписать так:

$$\frac{x^2}{2ph} + \frac{y^2}{2qh} = 1; \quad z = h, \quad (15)$$

т. е. в сечении поверхности плоскостями, параллельными плоскости XOY получаются эллипсы, полуоси которых возрастают по мере увеличения h .

2. Найдем линию пересечения поверхности (12) с плоскостью YOZ , уравнение которой $x = 0$. Совместное решение этих уравнений показывает, что в сечении поверхности (12) с плоскостью YOZ получится парабола (с осью симметрии OZ): $y^2 = 2qz$, $x = 0$.

3. Аналогично в сечении поверхности (12) плоскостью $y = 0$ получается также парабола $x^2 = 2pz$, $y = 0$ (с осью симметрии OZ).

Полученные линии пересечения позволяют изобразить поверхность

$$\frac{x^2}{p} + \frac{y^2}{q} = 2z, \quad (12)$$

которая называется *эллиптическим параболоидом* (рис. 104).

Рассмотренный ранее параболоид вращения является частным случаем эллиптического параболоида (12).

Гиперболический параболоид. Исследуем поверхность, заданную уравнением

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z, \quad (16)$$

где p и q имеют одинаковые знаки. Будем рассматривать случаи, когда $p > 0$ и $q > 0$.

1. Рассмотрим сечение поверхности (16) плоскостью XOZ : $y = 0$. В сечении плоскостью XOZ получается парабола $x^2 = 2pz$; $y = 0$, направленная вверх, с вершиной в начале координат и с осью симметрии OZ .

При пересечении поверхности плоскостями, параллельными плоскости XOZ ($y = h$), будут получаться параболы

$$x^2 = 2p\left(z + \frac{h^2}{2q}\right); \quad y = h,$$

«ветви» которых также направлены вверх.

2. Рассмотрим сечение поверхности (16) плоскостью YOZ ($x = 0$). В этом случае получится парабола $y^2 = -2qz$; $x = 0$ с вершиной в начале координат, с осью симметрии OZ , с «ветвями», направленными вниз, и как бы «сядущая верхом» на параболу $x^2 = 2pz$.

При пересечении поверхности плоскостями, параллельными плоскости YOZ ($x = h$), получаются параболы, уравнения которых будут:

$$y^2 = -2q\left(z - \frac{h^2}{2p}\right); \quad x = h.$$

Эти параболы «ветвями» направлены вниз и вершины их лежат на параболе

$$x^2 = 2pz; \quad y = 0.$$

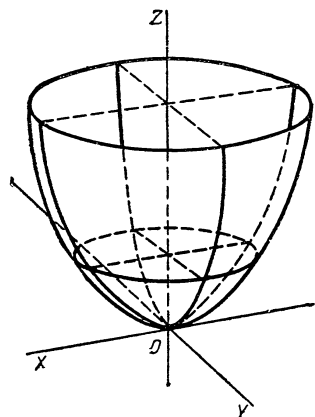


Рис. 104.

3. Рассмотрим сечения поверхности (16) плоскостями $z=h$, параллельными плоскости XOY .

При $h=0$ в плоскости $z=h$ получится линия пересечения $\frac{x^2}{p} - \frac{y^2}{q} = 0$, распадающаяся на пару пересекающихся прямых:

$$\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = 0; \quad \frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} = 0.$$

При $h>0$ в сечении будут получаться гиперболы

$$\frac{x^2}{2ph} - \frac{y^2}{2qh} = 1; \quad z=h, \quad (17)$$

действительными осями которых будут прямые, параллельные оси OX .

При $h<0$ уравнения (17) есть уравнения гипербол, действительными осями которых будут прямые, параллельные оси OY .

Совокупность линий пересечения данной поверхности плоскостями, параллельным координатным плоскостям, позволяет определить вид поверхности, заданной уравнением $\frac{x^2}{p} - \frac{y^2}{q} = 2z$ (рис. 105). Эта седлообразная поверхность называется *гиперболическим параболоидом*.

Конус второго порядка. Исследуем уравнение поверхности

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0. \quad (18)$$

Так как левая часть уравнения (18) является однородной функцией относительно x, y, z , то согласно изложенному в § 27 уравнение (18) является уравнением конической поверхности с вершиной в начале координат.

1. Рассмотрим сечение этой поверхности плоскостями $z=h$, параллельными плоскости XOY .

При $h=0$ получаем уравнения

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0; \quad z=0,$$

откуда следует, что линия пересечения поверхности плоскостью $z=0$ вырождается в точку $(0,0,0)$.

При $h \geq 0$ в сечениях будем получать эллипсы:

$$\frac{x^2}{a^2 h^2} + \frac{y^2}{b^2 h^2} = 1; \quad z=h.$$

Полуоси этих эллипсов будут возрастать при увеличении абсолютной величины h .

2. В сечении конической поверхности плоскостью XOZ ($y=0$) получится линия

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 0; \quad y=0,$$

распадающаяся на пару пересекающихся прямых:

$$\frac{x}{a} + \frac{z}{c} = 0; \quad y=0,$$

$$\frac{x}{a} - \frac{z}{c} = 0; \quad y=0.$$

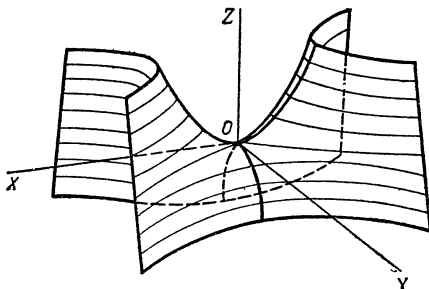


Рис. 105.

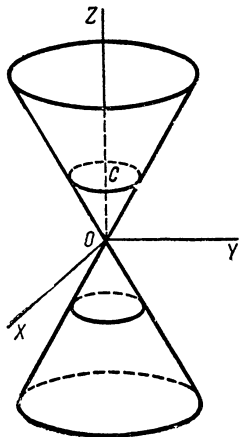


Рис. 106.

3. Аналогично в сечении плоскостью YOZ ($x=0$) получится также пара пересекающихся прямых

$$\frac{y}{b} + \frac{z}{c} = 0; \quad x = 0,$$

$$\frac{y}{b} - \frac{z}{c} = 0; \quad x = 0.$$

Полученные линии пересечения позволят изобразить поверхность (рис. 106)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0, \quad (19)$$

которая называется *конусом второго порядка*.

ГЛАВА III

ВВЕДЕНИЕ В АНАЛИЗ

Раздел 1

ФУНКЦИИ

§ 1. Числа

Как известно, числа появляются в результате счета и при измерении величин. Простейшим классом чисел является класс натуральных чисел. Натуральными числами называются числа $1, 2 = 1 + 1, 3 = 2 + 1, \dots$ Натуральные числа или, что то же, целые положительные числа, а также целые отрицательные числа $-1, -2, -3, \dots$ и число 0 , взятые вместе, образуют класс целых чисел. Дальнейшим расширением класса целых чисел является класс рациональных чисел. Рациональным числом называется частное двух целых чисел при условии, что делитель отличен от нуля.

Каждое рациональное число можно записать в виде конечной десятичной или бесконечной десятичной периодической дроби. Иррациональные числа записываются бесконечными десятичными непериодическими дробями. Рациональные и иррациональные числа, взятые вместе, образуют класс действительных чисел¹⁾, дальнейшим расширением которого является класс комплексных чисел. Комплексные числа — это числа вида $x + iy$, где x и y действительные числа, $i = \sqrt{-1}$.

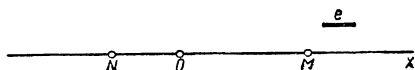


Рис. 107.

Основные сведения о комплексных числах даются в главе V.

Числовая ось. Прямая, на которой выбрано начало отсчета O , направление и масштаб (отрезок, принятый за единицу длины, который обозначим через e), называется числовой осью.

Между точками числовой оси (рис. 107) и действительными числами устанавливается взаимно однозначное соответствие по следующему закону: действительному числу a соответствует точка M с координатой a . Точка M лежит справа от начала O , если $a > 0$, слева от O , если $a < 0$, и, наконец, точка M совпадает с O , если $a = 0$.

При этом каждой точке N оси соответствует действительное число b — координата этой точки.

Имея в виду взаимно однозначное соответствие между действительными числами и точками числовой оси, вместо слова «число» часто говорят «точка» и наоборот.

¹⁾ В нашем курсе мы не будем рассматривать теорию действительных чисел.

Множество (совокупность) всех чисел x , удовлетворяющих условию $a < x < b$, где a и b — некоторые числа, называется *интервалом* (открытым интервалом). Интервал обозначается (a, b) . Множество всех чисел x , удовлетворяющих условию $a \leq x \leq b$, называется *отрезком* (замкнутым интервалом), или *сегментом*, оно обозначается $[a, b]$. Если к интервалу (a, b) присоединить числа a и b , то получится отрезок $[a, b]$. Точки a и b называются концами интервала (a, b) (отрезка $[a, b]$).

Множество всех чисел x , удовлетворяющих неравенствам $a < x \leq b$ или $a \leq x < b$, называются *полуинтервалами* и обозначаются соответственно $(a, b]$, $[a, b)$. Совокупность всех действительных чисел обозначают $(-\infty, +\infty)$. Совокупность чисел x , удовлетворяющих условию $x \leq a$,

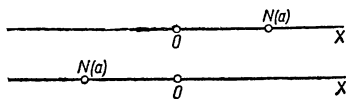


Рис. 108.

и совокупность чисел x , удовлетворяющих условию $x > a$, обозначают соответственно $(-\infty, a]$, $(a, +\infty)$, называют бесконечными *полуинтервалом* и *интервалом* и т. д.

Все рассмотренные выше множества чисел будем называть числовыми промежутками, или, короче, *промежутками*.

Окрестностью точки на числовой оси называется интервал (открытый интервал), содержащий эту точку. Здесь имеется в виду любой интервал, как симметричный, так и не симметричный относительно рассматриваемой точки.

Абсолютная величина действительного числа. *Абсолютная величина*, или *модуль*, действительного числа a обозначается $|a|$ и определяется следующим образом:

$$|a| = \begin{cases} a, & \text{если } a \geq 0, \\ -a, & \text{если } a < 0. \end{cases}$$

Например: $|-5,3| = 5,3$, $|+0,35| = 0,35$, $|-1,25| = 1,25$, $|0| = 0$.

На числовой оси $|a|$ есть длина отрезка ON , где N точка с координатой a (рис. 108).

Из определения абсолютной величины непосредственно следует, что

$$\begin{aligned} |a| &= |-a|, \\ a &\leq |a|, \\ -a &\leq |a|. \end{aligned}$$

Отметим также следующие свойства абсолютных величин:

1) Неравенство $|a| < M$ равносильно двум неравенствам: $-M < a < M$.

Действительно, $a \leq |a| < M$, следовательно, $a < M$, $-a \leq |a| < M$, следовательно, $-a < M$, т. е. $a > -M$. И окончательно, имеем

$$-M < a < M.$$

Обратно, если $-M < a < M$, то $|a| < M$. Действительно, если $a \geq 0$, то $|a| = a$ и, учитывая правое неравенство, получим $|a| < M$; если $a < 0$, то $|a| = -a$ и, умножая левое неравенство на -1 , получим:

$$|a| < M.$$

Аналогично доказывается, что запись $|a| \leq M$ равносильна записи: $-M \leq a \leq M$.

Неравенство $|a| < M$ ($|a| \leq M$) означает, что число a принадлежит интервалу $(-M, M)$ (отрезку $[-M, M]$).

2) Абсолютная величина суммы двух чисел не больше суммы абсолютных величин слагаемых: $|a + b| \leq |a| + |b|$.

Действительно, пусть $a + b \geq 0$.

Как известно, $a \leq |a|$, $b \leq |b|$, складывая эти неравенства, получим $a + b \leq |a| + |b|$. Но $|a + b| = a + b$ и, следовательно, $|a + b| \leq |a| + |b|$.

Пусть теперь $a + b < 0$. Так как $-a \leq |a|$, $-b \leq |b|$, то, складывая эти неравенства, имеем $-(a + b) \leq |a| + |b|$. Но $|a + b| = -(a + b)$, поэтому окончательно, получим:

$$|a + b| \leq |a| + |b|.$$

Это свойство легко распространяется на случай любого числа слагаемых. Так как $|a+b+c+\dots+k+l|\leq|a|+|b+c+\dots+k+l|\leq|a|+|b|+|c+\dots+k+l|$ и т. д., то

$$|a+b+c+\dots+k+l|\leq|a|+|b|+|c|+\dots+|k|+|l|.$$

3) Абсолютная величина разности двух чисел не меньше разности абсолютных величин этих чисел: $|a-b|\geq||a|-|b||$.

Действительно, рассмотрим равенство $a=(a-b)+b$.
По свойству 2)

$$|a|=|(a-b)+b|\leq|a-b|+|b|,$$

откуда следует, что $|a-b|\geq|a|-|b|$

$$4) |a-b|\geq||a|-|b||.$$

По свойству 3)

$$|a-b|\geq|a|-|b| \quad (a)$$

и

$$|a-b|=|b-a|\geq|b|-|a|. \quad (б)$$

По определению абсолютная величина разности $|a|-|b|$ равна или $|a|-|b|$ или $|b|-|a|$. Учитывая соотношения (a) и (б), имеем:

$$||a|-|b||\leq|a-b|.$$

$$5) |ab|=|a|\cdot|b|.$$

$$6) \left|\frac{a}{b}\right|=\frac{|a|}{|b|}.$$

Свойства 5 и 6 непосредственно следуют из определения абсолютной величины числа.

§ 2. Функции и их графики

При рассмотрении количественных соотношений реального мира приходится сталкиваться с численными значениями различных физических величин, как, например, времени, температуры, объема, длины, давления, плотности и пр.

В зависимости от рассматриваемых условий одни из величин имеют постоянные численные значения, у других эти значения меняются. Такие величины называются, соответственно *постоянными* и *переменными в данных условиях*¹⁾. Существуют величины, сохраняющие постоянные численные значения независимо от условий задачи, например атомная масса элемента, величина прямого угла и т. д. Такие величины называют *абсолютными постоянными*.

Внимательное изучение окружающих нас явлений показывает, что физические величины изменяются не независимо друг от друга, а так, что изменение численных значений одних из этих величин влечет за собой изменение значений других.

Математический анализ как раз и занимается изучением количественных соотношений совместно изменяющихся величин, отвлекаясь от их конкретного физического смысла.

В связи с этим в математическом анализе под *переменной величиной* (просто *переменной*) подразумевается отвлеченная, числовая переменная. Переменная (обозначим ее буквой x) считается заданной, если указано множество значений, которые она может принимать. *Постоянную величину* (просто *постоянную*) удобно рассматривать как такой частный случай переменной, когда множество принимаемых ею значений состоит из одного числа.

¹⁾ Величины, принимающие постоянные численные значения только в определенных условиях, называют *параметрами*. Так, например, в уравнении прямой $y=kx+b$ числа k и b являются параметрами.

Одним из основных понятий математического анализа является понятие функции.

Переменная y называется *функцией* от другой переменной x , если каждому значению x из некоторого множества по определенному правилу или закону ставится в соответствие одно или несколько значений переменной y .

При этом переменная x называется *независимой переменной*, или *аргументом*, а переменная y — *зависимой переменной*, или *функцией*.

Множество значений, принимаемых переменной x , называется *областью определения* функции, а множество значений, принимаемых переменной y , называется *областью изменения* функции. В дальнейшем в основном будут рассматриваться множества, состоящие только из действительных чисел как для аргумента x , так и для функции y .

Если каждому значению x из области определения функции соответствует одно значение y , то функция называется *однозначной*. В противном случае — *многозначной*.

Приведем несколько примеров.

Очевидно, что если переменная x может принимать любое числовое значение, то переменные $y=x^2$, $y=\cos x$ и другие будут функциями переменной x . Областью определения каждой из этих функций будет по условию множество всех действительных чисел. Область изменения функции $y=x^2$ есть множество всех неотрицательных действительных чисел, а область изменения функции $y=\cos x$ состоит из действительных чисел, удовлетворяющих неравенствам $-1 \leq y \leq 1$.

Путь s , пройденный телом, падающим в пустоте, за время t , есть функция времени t , так как между переменными s и t имеет место соотношение

$$s = \frac{gt^2}{2}$$

(где g — ускорение силы тяжести).

Область определения этой функции — множество значений t от $t=0$ до $t=t_1$, где t_1 — время, необходимое для падения тела. Область изменения функции дается неравенствами $0 \leq s \leq s_1$, где $s_1 = \frac{gt_1^2}{2}$. Каждому значению t из области определения соответствует определенное значение s .

Если каждому значению переменной x отрезка $[-2, 2]$ поставить в соответствие те значения y , для которых $\frac{x^2}{4} + \frac{y^2}{9} = 1$, т. е. поставить в соответствие $y = \pm 3 \sqrt{1 - \frac{x^2}{4}}$, мы получим многозначную (двузначную) функцию x . Областью определения этой функции является отрезок $[-2, 2]$, а областью изменения — отрезок $[-3, 3]$. В дальнейшем, если это специально не оговаривается, будут рассматриваться только однозначные функции.

Для того чтобы обозначить тот факт, что переменная y есть функция переменной x , обычно пишут:

$$y = f(x), y = \varphi(x), y = F(x) \text{ и т. д.}$$

Буквы f , φ , F и т. д. используются здесь для обозначения закона соответствия между аргументом x и функцией y . Иногда и закон соответствия и результат соответствия обозначают одной и той же буквой, например,

$$y = y(x).$$

Запись $f(x_0)$ обозначает, что берется частное значение функции $f(x)$ при $x = x_0$.

Например, если $y = x^2 - 2x + 3$, где $-\infty < x < +\infty$, то $f(x) =$

$=x^2-2x+3$. При $x=2$ мы имеем частное значение функции: $f(2)=2^2-2\cdot 2+3=3$.

Для наглядного представления изменения функции в зависимости от изменения аргумента часто пользуются графическим изображением функции.

Рассмотрим на плоскости прямоугольную декартову систему координат XOY . Значение аргумента x функции $y=f(x)$ будем изображать точкой с абсциссой x на оси OX . Тогда область определения функции $y=f(x)$ будет изображаться некоторым множеством точек оси OX .

Для каждой точки x_0 этого множества построим на плоскости точку M_0 с координатами (x_0, y_0) , где $y_0=f(x_0)$. Совокупность всех таких точек составляет график функции $y=f(x)$ и служит для ее геометрической иллюстрации (см. рис. 109). Таким образом, *графиком* функции $y=f(x)$ на плоскости XOY называется геометрическое место точек плоскости с координатами $x, f(x)$, где переменная x принимает все значения из области определения функции.

Предположим для простоты, что область определения функции $y=f(x)$ изображается некоторым промежутком оси OX . График такой функции представляет собой некоторую кривую. Ординаты точек этой кривой для различных значений x равны соответствующим значениям функции $f(x)$.

Построение графика функции можно производить по точкам¹⁾. Для этого в данном промежутке берется ряд близких между собой значений x (например, $x_1, x_2, \dots, x_n$), вычисляются соответствующие им значения функции y ($y_1, y_2, \dots, y_n$), затем строятся на плоскости точки $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ и через них плавной проводится кривая. Эта кривая приближенно изображает график данной функции. Чем чаще взяты точки $x_1, x_2, \dots, x_n$, тем, вообще говоря, точнее построенная кривая воспроизводит график функции.

Способы задания функций. Согласно определению, функция считается заданной, если задано множество значений, принимаемых независимой переменной (область определения функции), и задано правило, устанавливающее соответствие между значениями независимой и зависимой переменных. При этом никаких ограничений на характер этого соответствия не накладывается.

В математическом анализе функция преимущественно задается с помощью формулы, указывающей, какие действия надо произвести над аргументом, чтобы получить соответствующие значения функции. Такая формула называется *аналитическим выражением* функции, а способ задания функции с помощью формулы называется *аналитическим способом* задания функции.

Так, например, поверхность шара S есть функция его радиуса R . Ее аналитическое выражение:

$$S=4\pi R^2, \text{ где } 0 < R < +\infty.$$

Формула $y=x+\sqrt{x^2+4}$ определяет функцию y для всякого значения аргумента x .

В частности, функция может быть задана и с помощью нескольких формул (разные формулы на разных частях области определения). Например,

$$f(x)=\begin{cases} x^2, & -3 < x < 0, \\ 0, & 0 \leq x \leq 1, \\ (x-1)^2, & 1 < x < 10. \end{cases}$$

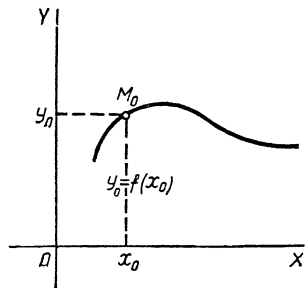


Рис. 109.

¹⁾ Ниже будут указаны более точные и совершенные методы построения графиков функций.

Область определения этой функции — множество действительных чисел x , принадлежащих интервалу $-3 < x < 10$. Каждому значению x из этого интервала поставлено в соответствие определенное значение функции. При этом $f(-2) = (-2)^2 = 4$, $f\left(\frac{1}{2}\right) = 0$, $f(5) = (5-1)^2 = 16$ и т. д.

При исследовании явлений природы иногда приходится встречаться с переменными величинами, зависимость между которыми устанавливается опытным путем. В таких случаях на основании опытных данных составляются таблицы, в которых содержатся значения функций, соответствующие различным частным значениям аргумента. Этот способ задания функций называется *табличным*. Примеры табличного задания функций можно найти в любом техническом справочнике.

В ряде случаев, особенно при использовании самопишущих приборов, функция задается непосредственно графиком (*графический способ задания функции*). Например, кривые, появляющиеся на лентах осциллографа, могут представлять собой графики различных физических величин как функций времени. С этих графиков для любого момента времени можно «снять» (приблизленно) значения функций.

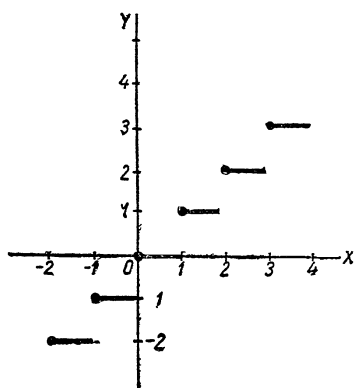


Рис. 110.

Указанные выше три способа задания функции (аналитический, табличный и графический), являясь наиболее употребительными, не исчерпывают всех возможных способов.

В частности, можно задать функцию, описав словами, какие значения она принимает для различных значений аргумента.

Например, функция $E(x)$ («антье икс»), ее иногда обозначают также символом $[x]$, определяется как наибольшее целое число, не превосходящее x . При этом $E(1) = 1$, $E(3,1) = 3$, $E(0,7) = 0$, $E(-0,5) = -1$.

Другой пример — функция Дирихле, определяемая следующим образом: $f(x)$ равна нулю для всех иррациональных значений x и равна единице для всех рациональных значений x .

Примечание. Задание функции каким-либо способом не исключает возможности ее задания и другими способами. Функцию $E(x)$ и функцию Дирихле также можно задать аналитически. Ввиду сложности этого задания им обычно не пользуются. График функции $E(x)$ см. на рис. 110.

Остановимся несколько подробнее на аналитическом способе задания функции.

Начнем с примера. Формула $S = 4\pi R^2$ определяет площадь поверхности шара S как функцию радиуса шара R . По смыслу задачи эта функция определена в области всех положительных чисел R ($0 < R < +\infty$).

Отвлекаясь от содержания задачи, рассмотрим формулу $S = 4\pi R^2$. Эта формула каждому значению переменной R , $-\infty < R < +\infty$, ставит в соответствие определенное значение переменной S , т. е. определяет S как функцию R в области всех действительных значений R .

Таким образом, область определения функции S в формуле для вычисления площади поверхности оказалась лишь частью той естественной области, в которой формула $S = 4\pi R^2$ имеет смысл.

С такими явлениями приходится часто сталкиваться при решении конкретных смысловых задач.

В дальнейшем нам неоднократно придется рассматривать аналитически заданные функции в их естественных областях определения.

При этом *естественной областью определения* аналитически заданной функции мы будем называть совокупность всех действительных значений независимой переменной, для которых функция принимает действительные значения.

Так, например, для функции $y = \frac{1}{4+x^2}$ естественной областью определения будет множество всех действительных чисел; для функции $y = \frac{1}{1-x^2}$ естественная область определения задается неравенствами $-\infty < x < -1$, $-1 < x < 1$, $1 < x < +\infty$; функция $y = \sqrt{x}$ определена для всех неотрицательных чисел, функция $y = \sqrt{1-x^2}$ определена на отрезке $[-1, 1]$ и т. д.¹⁾

Условимся в тех случаях, когда функция задана аналитически и не указана дополнительно к формуле или смыслом рассматриваемой задачи ее область определения, подразумевать естественную область определения.

Понятие обратной функции. Пусть функция $y=f(x)$ определена в некоторой области. Возьмем какое-нибудь значение y_0 из множества значений этой функции. В области определения функции найдется одно или несколько²⁾ значений аргумента x_0 , для которых функция принимает значение y_0 , т. е. $f(x_0)=y_0$. Поставим y_0 в соответствие все эти значения x_0 .

При этом каждому значению переменной y из множества значений функции $y=f(x)$ ставится в соответствие одно или несколько значений x , т. е. на множестве значений функции $y=f(x)$ задается однозначная или многозначная функция $x=\varphi(y)$. Она называется *обратной* для функции $y=f(x)$.

Областью определения обратной функции является область изменения данной функции.

Например, функция $y=x^3$, $-\infty < x < +\infty$, имеет областью изменения все действительные значения переменной y . Для любого значения y из промежутка $-\infty < y < +\infty$ существует одно значение x такое, что $x^3=y$. Это значение x равно $\sqrt[3]{y}$.

Таким образом, $x=\sqrt[3]{y}$ есть однозначная обратная функция для функции $y=x^3$. Ее область определения $-\infty < y < +\infty$, а множество значений: $-\infty < x < +\infty$.

Функция $y=x^2$ определена в промежутке $-\infty < x < +\infty$. Ее область изменения: $0 \leq y < +\infty$. Для каждого значения y из этой области существуют два значения переменной x : $x=+\sqrt{y}$ и $x=-\sqrt{y}$, таких, что $x^2=y$. Следовательно, функция, обратная к функции $y=x^2$, будет двухзначной. Ее можно записать в виде: $x=\pm\sqrt{y}$. Область определения такой функции: $0 \leq y < +\infty$, область изменения: $-\infty < x < +\infty$.

Рассмотрим теперь вопрос о графике обратной функции.

Данная функция $y=f(x)$ и ее обратная функция $x=\varphi(y)$ выражают, по существу, одну и ту же связь между переменными x и y , только в первом случае мы рассматриваем x как аргумент, y как функцию, а во втором случае — наоборот. Поэтому график обратной функции $x=\varphi(y)$ совпадает с графиком данной функции $y=f(x)$ (рис. 111).

Если у обратной функции, так же как и у данной, аргумент обозначить через x , а значения функции через y , то она запишется в виде $y=\varphi(x)$.

¹⁾ Не всякое аналитическое выражение задает функцию. Например, выражение $\sqrt{1-x^2} + \sqrt{x^2-9}$ в области действительных чисел никакой функции не определяет, так как не существует действительных значений x , для которых оно принимало бы действительные значения.

²⁾ Или даже бесконечно много.

Функции $x = \varphi(y)$, $y = \varphi(x)$ различаются только обозначениями переменных. Поэтому, чтобы из графика функции $x = \varphi(y)$ [или, что то же, функции $y = f(x)$] получить график функции $y = \varphi(x)$, достаточно поменять ролями оси OX и OY , т. е. повернуть плоскость чертежа вокруг биссектрисы первого координатного угла на 180° .

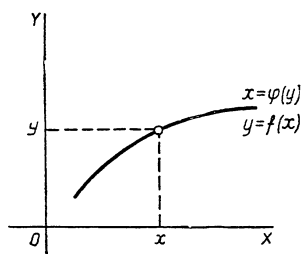


Рис. 111.

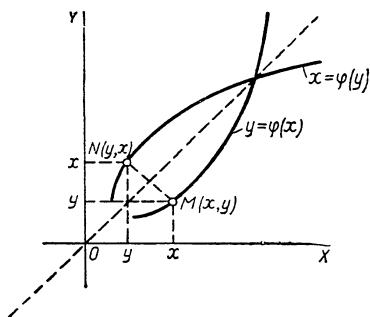


Рис. 112.

Таким образом, график обратной функции $y = \varphi(x)$ симметричен графику данной функции $y = f(x)$ относительно биссектрисы первого координатного угла (см. рис. 112, 113, 114).

Сложная функция. Пусть даны две функции $u = \varphi(x)$ и $y = F(u)$, причем множество значений первой функции входит в область определения второй. Тогда каждому значению x из области определения функции $u = \varphi(x)$ соответствует определенное значение переменной u , а этому

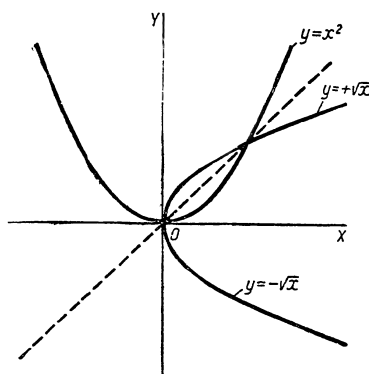


Рис. 113.

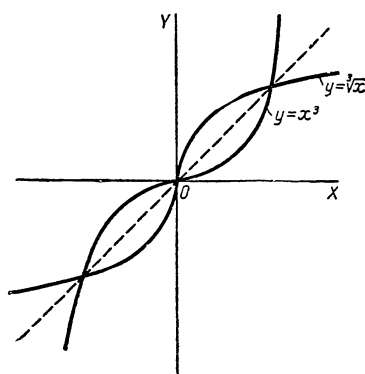


Рис. 114.

значению u функция $y = F(u)$ ставит в соответствие определенное значение переменной y , т. е. переменная y является функцией от x :

$$y = F(\varphi(x)).$$

Полученная функция от функции называется *сложной функцией*¹⁾ переменной x . Функцию $\varphi(x)$ называют при этом *промежуточной переменной*.

Например, если $u = x^2 + 1$, $y = \sqrt{u}$, то для всех значений x определена сложная функция $y = \sqrt{x^2 + 1}$. Здесь $x^2 + 1$ — промежуточная переменная.

Сложными функциями будут $2^{\sin x}$, $\cos(2x - 3)$, $\lg \frac{1}{x}$ и т. д.

¹⁾ Сложную функцию $y = F(\varphi(x))$ называют часто *суперпозицией* (наложением) двух функции: *внутренней* $u = \varphi(x)$ и *внешней* $y = F(u)$.

Неявная функция. Пусть дано уравнение, связывающее две переменные x и y . Если все его члены перенести в левую часть, оно запишется в виде:

$$F(x, y) = 0.$$

Может быть, это уравнение не удовлетворяется ни одной парой действительных чисел x и y . Таким, например, будет уравнение

$$x^2 + y^2 + 1 = 0.$$

Может быть, уравнение удовлетворяется лишь одной парой действительных чисел, как, например, уравнение $(x-1)^2 + (y-2)^2 = 0$, удовлетворяющееся только парой чисел $x=1, y=2$. И, наконец, может быть, существуют различные пары действительных чисел (x, y) , удовлетворяющие данному уравнению. Например, уравнению $2x - y + 1 = 0$ удовлетворяют пары чисел $(x, 1 + 2x)$, где x — любое действительное число.

В этом последнем случае на уравнение $F(x, y) = 0$ можно смотреть как на способ задания переменной y как функции от x . А именно, с по-

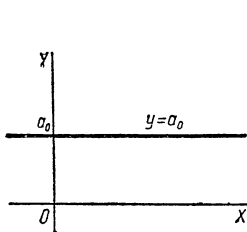


Рис. 115.

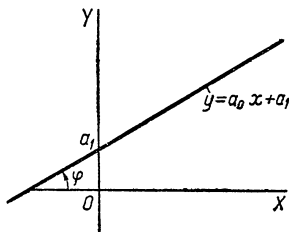


Рис. 116.

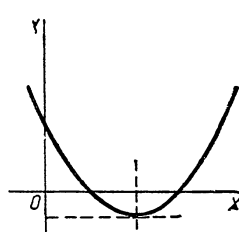


Рис. 117.

мощью этого уравнения каждому значению x из некоторого множества ставится в соответствие одно или несколько значений переменной y [таких, что для них $F(x, y) \equiv 0$].

Функция от x , определяемая уравнением $F(x, y) = 0$, называется *неявно заданной* этим уравнением. Так, уравнение $2x - y + 1 = 0$ неявно задает y как функцию от x .

В ряде случаев, разрешив уравнение $F(x, y) = 0$ относительно y , удастся получить явное задание той же функции. Например, уравнение $2x - y + 1 = 0$ неявно задает функцию $y = 1 + 2x$; уравнение $x^2 + y^2 - 1 = 0$ неявно задает двухзначную функцию $y = \pm \sqrt{1 - x^2}$ и т. д.

Элементарные функции и их графики. Рассмотрим некоторые классы функций, наиболее часто встречающихся в математическом анализе.

1. Многочлены, или целые рациональные функции. *Многочленом, целой рациональной функцией* или *полиномом* относительно переменной x называется функция вида

$$y = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n, \quad (1)$$

где $a_0, a_1, \dots, a_n$ — постоянные, называемые *коэффициентами* многочлена. Старшая входящая в многочлен степень переменной x называется *степенью* многочлена. Если $a_0 \neq 0$, то степень многочлена (1) равна n .

Область определения многочлена — все действительные значения переменной x .

При $n=0$ имеем многочлен нулевой степени: $y = a_0$, где a_0 — число. Эта функция сохраняет одно и то же постоянное значение для всех x . Графиком служит прямая, параллельная оси OX (рис. 115).

При $n=1$ имеем многочлен первой степени: $y = a_0 x + a_1$, его называют *линейной функцией* относительно x . Графиком является прямая, наклоненная к оси OX под углом φ , для которого $\operatorname{tg} \varphi = a_0$, и отсекающая на оси OY отрезок, величина которого a_1 (рис. 116).

При $n=2$ имеем многочлен второй степени: $y=a_0x^2+a_1x+a_2$, который называют *квадратичной функцией* относительно x . Графиком квадратичной функции будет парабола с осью, параллельной оси $OY^1)$ (рис. 117 для $a_0>0$).

2. Дробно-рациональные функции. *Дробно-рациональной функцией* относительно x называется функция, которую можно представить в виде отношения двух многочленов относительно x , т. е. в виде

$$y = \frac{a_0x^n + a_1x^{n-1} + \dots + a_n}{b_0x^m + b_1x^{m-1} + \dots + b_m},$$

где $a_0, \dots, a_n, b_0, \dots, b_m$ — постоянные.

Она определена для всех значений x , кроме тех, для которых знаменатель равен нулю.

Примерами дробно-рациональных функций являются

$$y = \frac{x-1}{x^2+1}, \quad y = \frac{x^3-3x^2+2}{x^3-9}, \quad y = \frac{1}{x^2} - \frac{2x+1}{3x-1}$$

и т. д.

Функция $y = \frac{a_0x+a_1}{b_0x+b_1}$, где a_0, a_1, b_0, b_1 — постоянные, причем $b_0 \neq 0$ и $a_0b_1 \neq a_1b_0$, называется *дробно-линейной функцией* от x . Это частный

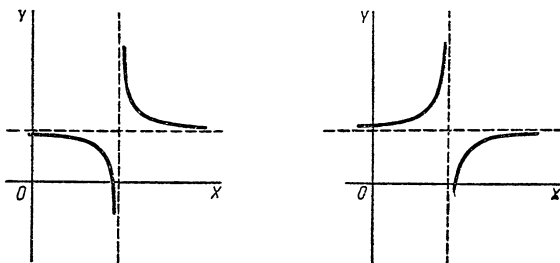


Рис. 118.

случай дробно-рациональной функции. Графиком такой функции является гипербола²⁾ (рис. 118).

3. Алгебраические функции. *Алгебраической* называют функцию, полученную из аргумента и действительных чисел с помощью алгебраических операций (сложение, вычитание, умножение, деление, возведение в степень и извлечение корня), а также любую неявную функцию, связанную с аргументом алгебраическим уравнением какой-нибудь степени.

Из определения следует, что всякая рациональная относительно x функция (целая или дробная) является алгебраической.

Алгебраические функции, не являющиеся рациональными, называют *иррациональными*.

Примеры алгебраических функций:

$$y = \sqrt{x-1} + \sqrt[3]{x+3}; \quad y = \left(\frac{x^2-2}{x+\sqrt{x}} \right)^5; \quad y = 2x-1;$$

функция $y=f(x)$, заданная неявно уравнением $x^2y^2-2xy+3=0$ и т. д.

Частным случаем алгебраической функции является *степенная* функция $y=x^\alpha$, где α — рациональное число ($\alpha = \frac{p}{q}$, где p и q — целые числа).

Если q — нечетное число, а p — любое число, то функция $y=x^\alpha$ определена для всех действительных значений x (кроме $x=0$ при $\alpha<0$).

¹⁾ Построение графика квадратичной функции дается в § 3.

²⁾ Построение графика дробно-линейной функции см. в § 3.

Если q — четное число, а p — нечетное число, то функция $y = x^{\frac{q}{p}}$ определена для $x \geq 0$ при $\alpha > 0$ и $x > 0$ при $\alpha < 0$.

Графики некоторых степенных функций см. на рис. 119—124.

4. Трансцендентные функции. Трансцендентными функциями называются все функции, не являющиеся алгебраическими. К ним относятся,

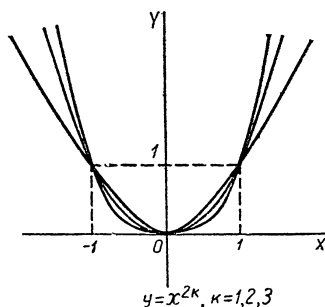


Рис. 119.

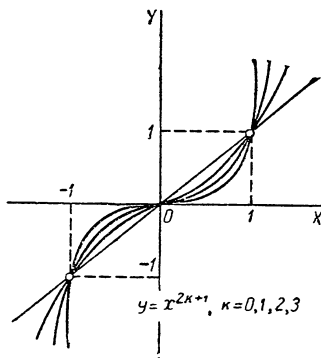


Рис. 120.

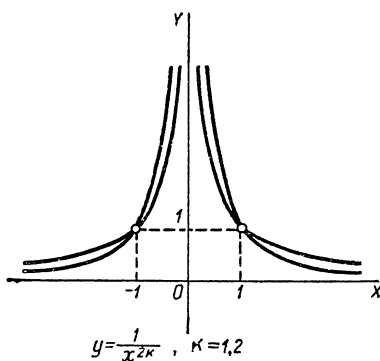


Рис. 121.

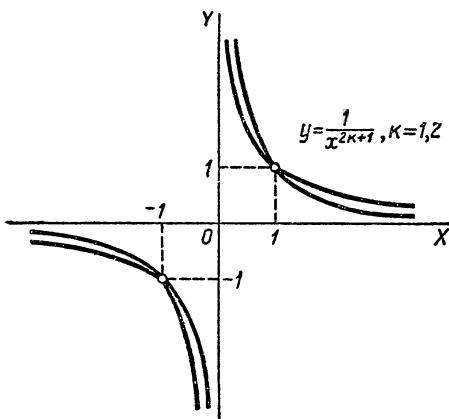


Рис. 122.

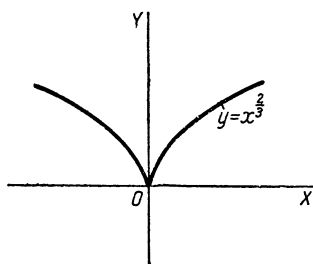


Рис. 123.

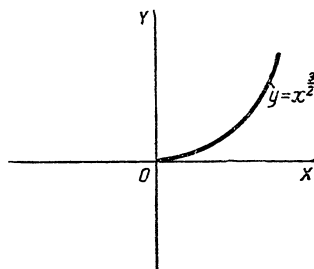


Рис. 124.

в частности, степенные функции $y = x^{\alpha}$, где α — любое иррациональное число, показательные, логарифмические, тригонометрические и обратные тригонометрические функции.

Рассмотрим графики некоторых из этих функций.

а) Показательная функция, т. е. функция вида $y = a^x$, где a — любое положительное число, отличное от единицы, определена для всех значений x .

Графики показательных функций для случая $a > 1$ и $0 < a < 1$ даны на рис. 125, 126.

б) *Логарифмическая* функция, т. е. функция вида $y = \log_a x$, где $a > 0$, $a \neq 1$, определяется как функция, обратная к показательной $y = a^x$.

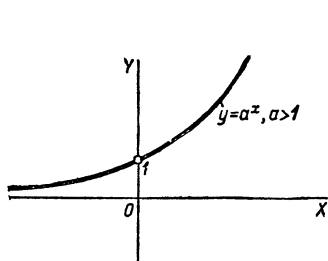


Рис. 125.

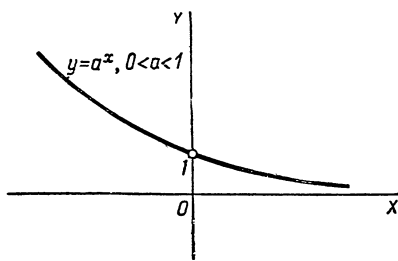


Рис. 126.

Область определения логарифмической функции — множество всех положительных значений аргумента x . График симметричен графику показательной функции $y = a^x$ относительно биссектрисы первого координатного угла (см. рис. 127, 128).

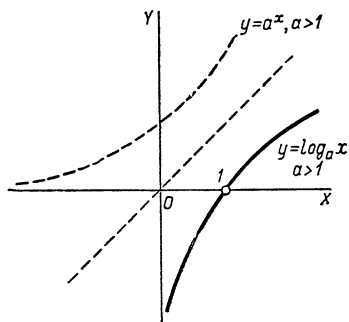


Рис. 127.

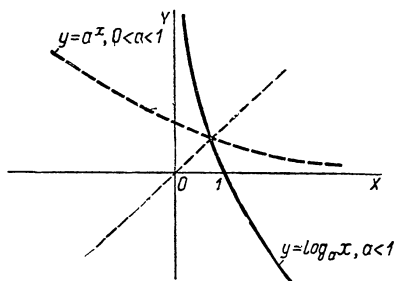


Рис. 128.

в) *Тригонометрические* функции

$$y = \sin x, \quad y = \cos x, \quad y = \operatorname{tg} x, \quad y = \operatorname{ctg} x, \quad y = \sec x, \quad y = \operatorname{cosec} x.$$

Здесь, как это принято в математическом анализе, каждое значение аргумента x есть отвлеченное число, представляющее собой радианную меру некоторого угла.

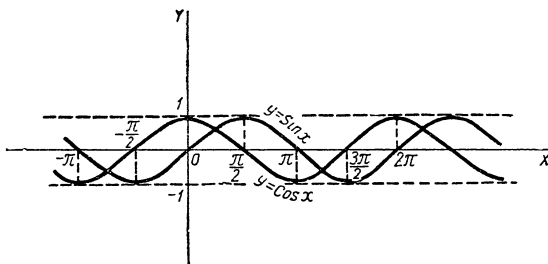


Рис. 129.

Функции $y = \sin x$ и $y = \cos x$ имеют областью определения все значения переменной x . Множеством значений каждой из этих функций будет промежуток: $-1 \leq y \leq 1$.

Графики функции $y = \sin x$ и $y = \cos x$ (их называют соответственно *синусоидой* и *косинусоидой*) см. на рис. 129.

Функция $y = \operatorname{tg} x$ определена для всех значений x , кроме $x = \frac{\pi}{2}(2k+1)$, $k=0, \pm 1, \pm 2, \dots$. Множество значений: $-\infty < y < +\infty$.

График этой функции (*тангенсоида*) дан на рис. 130.

Функция $y = \operatorname{ctg} x$ определена для всех значений x , кроме $x = k\pi$, $k=0, \pm 1, \pm 2, \dots$. Множество ее значений: $-\infty < y < +\infty$.

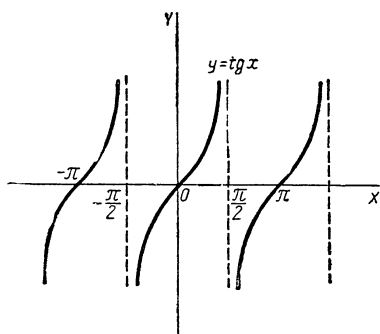


Рис. 130.

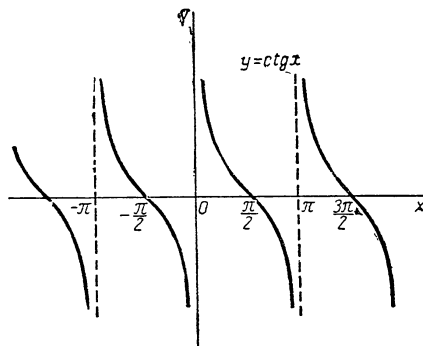


Рис. 131.

График функции $y = \operatorname{ctg} x$ (*котангенсоида*) приводится на рис. 131.

г) *Обратные тригонометрические функции.*

$y = \operatorname{Arcsin} x$ — бесконечнозначная функция, обратная для функции $y = \sin x$. Область определения: $-1 \leq x \leq 1$, область изменения: $-\infty < y < +\infty$.

Если каждому значению x , $-1 \leq x \leq 1$, поставить в соответствие значение y бесконечнозначной функции $y = \operatorname{Arcsin} x$, удовлетворяющее неравенствам: $-\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$, то

получится однозначная функция. Ее обозначают $y = \arcsin x$ и называют *главной ветвью* функции $y = \operatorname{Arcsin} x$. На рис. 132 ее график выделен жирной линией.

Функция $y = \operatorname{Arccos} x$ — бесконечнозначная функция, обратная для функции $y = \cos x$. Область определения: $-1 \leq x \leq 1$, область изменения: $-\infty < y < +\infty$. Поставив в соответствие каждому значению x из области определения функции $y = \operatorname{Arccos} x$ то из значений этой функции, которое лежит в пределах $0 \leq y \leq \pi$, получим однозначную функцию, называемую *главной ветвью* функции $y = \operatorname{Arccos} x$. Ее обозначают $y = \arccos x$. График функции $y = \arccos x$ на рис. 133 выделен жирной линией.

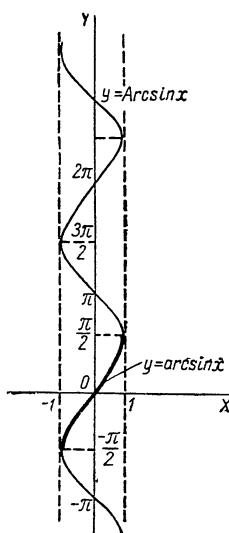


Рис. 132.

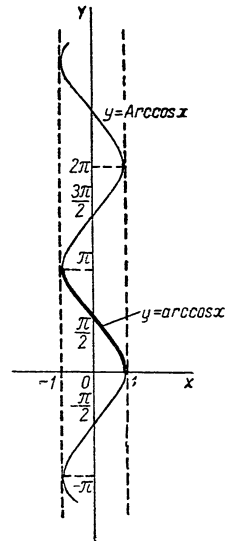


Рис. 133.

На рис. 134, 135 изображены графики бесконечнозначных функций $y = \operatorname{Arctg} x$ и $y = \operatorname{Arcctg} x$, определяемых как обратные функции для функций $y = \operatorname{tg} x$ и $y = \operatorname{ctg} x$ соответственно.

Области определения этих функций: $-\infty < x < +\infty$, области изменения: $-\infty < y < +\infty$.

Если каждому значению x , $-\infty < x < +\infty$, поставить в соответствие значение y функции $y = \text{Arctg } x$ ($y = \text{Arcctg } x$), удовлетворяющее неравенствам $-\frac{\pi}{2} < y < \frac{\pi}{2}$ ($0 < y < \pi$), то получится однозначная функция, называемая *главной ветвью* бесконечнозначной функции $y = \text{Arctg } x$ ($y = \text{Arcctg } x$) и обозначаемая $y = \text{arctg } x$ ($y = \text{arccctg } x$). График функции $y = \text{arctg } x$ ($y = \text{arccctg } x$) на рис. 134 (135) выделен жирной линией.

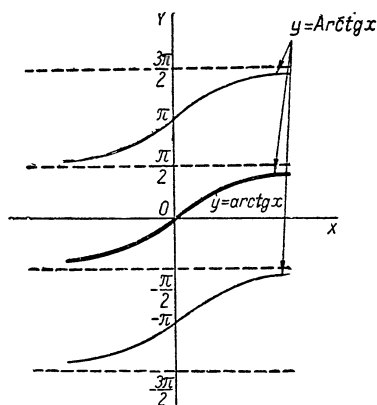


Рис. 134.

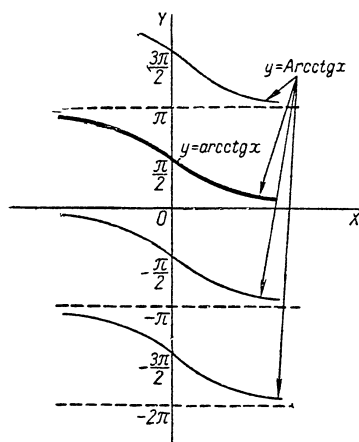


Рис. 135.

Будем называть *основными элементарными функциями* алгебраические функции, задаваемые явно, а также перечисленные выше трансцендентные функции ($y = x^a$, a — иррациональное число, $y = a^x$, $y = \log_a x$, $y = \sin x$, $y = \cos x$, ..., $y = \arcsin x$, и т. д.).

Функции, полученные из основных элементарных функций с помощью конечного числа алгебраических действий и применением конечного числа операций вычисления функции от функции, будем называть *элементарными функциями*.

§ 3. Некоторые замечания о построении графиков функций

Четные и нечетные функции. Построение графиков функций облегчается, если функция является четной или нечетной.

Функция $f(x)$, определенная на промежутке, симметричном относительно $x=0$, и, в частности, на всей числовой оси, называется *четной*, если для любого значения x этого промежутка $f(-x) = f(x)$, и *нечетной*, если $f(-x) = -f(x)$.

Например, функции $f(x) = x^2 + 2$, $\varphi(x) = \cos x$, $\psi(x) = 2^x + 2^{-x}$ — четные. Функции $f(x) = x^3 + 3x^3 + x$, $\varphi(x) = \text{tg } x$, $\psi(x) = 2^x - 2^{-x}$ — нечетные.

Заметим, что не всякая функция является четной или нечетной. Однако любую функцию, определенную в симметричном относительно $x=0$ промежутке, можно записать в виде суммы четной и нечетной функций:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}.$$

Из определения четной и нечетной функций непосредственно следует, что график четной функции симметричен относительно оси OY (рис. 136), а график нечетной функции симметричен относительно начала координат (рис. 137).

Поэтому для построения графиков четной или нечетной функций достаточно построить их графики для неотрицательных значений x . Затем,

в случае четной функции, построить график, симметричный полученному относительно оси OY , а в случае нечетной функции построить график, симметричный полученному относительно начала координат.

Сложение графиков. Для построения графика функции $y = f(x) + \varphi(x)$ при условии, что графики функций $y = f(x)$ и $y = \varphi(x)$ уже известны, можно воспользоваться сложением графиков, которое производится следующим образом. Возьмем какое-нибудь значение x из области опреде-

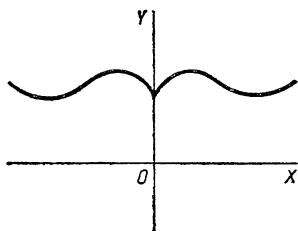


Рис. 136.

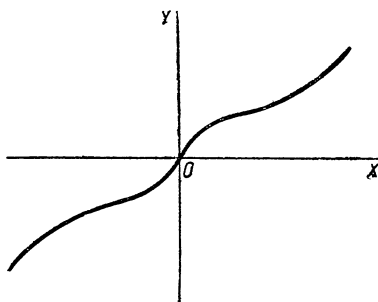


Рис. 137.

ления функций $f(x)$ и $\varphi(x)$ и произведем сложение направленных отрезков, величины которых равны значениям функций $f(x)$ и $\varphi(x)$ при рассматриваемом значении x . Взяв различные значения x , построим точки, принадлежащие графику суммы $f(x) + \varphi(x)$.

Так, например, можно построить график функции $y = x + \frac{1}{x}$ (см. рис. 138).

Деформация графиков. Рассмотрим построение графиков функций $f(x) + a$, $f(x + a)$, $af(x)$, $f(ax)$ при условии, что график функции $f(x)$ известен.

Задача 1. Зная график функции $f(x)$, построить график функции $f(x) + a$.

Решение. Сложив графики функций $y = f(x)$ и $y = a$, получим график $f(x) + a$. Нетрудно видеть, что для построения графика $f(x) + a$ достаточно сместить график функции $f(x)$ в вертикальном направлении на $|a|$ единиц масштаба — вверх при $a > 0$ и вниз при $a < 0$ или, что то же, оставляя график функции $f(x)$ на месте, сместить ось OX в вертикальном направлении на $|a|$ единиц масштаба — вниз, если $a > 0$, и вверх, если $a < 0$.

Задача 2. Построить график функции $f(x + a)$, если график функции $f(x)$ известен.

Решение. Функция $f(x + a)$ в точке $x - a$ принимает такое же значение, как и функция $f(x)$ в точке x . Поэтому для построения графика функции $f(x + a)$ нужно график функции $f(x)$ сместить в горизонтальном направлении на $|a|$ единиц масштаба — вправо при $a < 0$, влево при $a > 0$. Или, если оставить график функции $f(x)$ на месте, то следует сместить ось OY по горизонтали на $|a|$ единиц масштаба — влево при $a < 0$ и вправо при $a > 0$. Таким образом построен график функции $y = \log_a(x - 1) + 2$ ($a > 1$) (см. рис. 139).

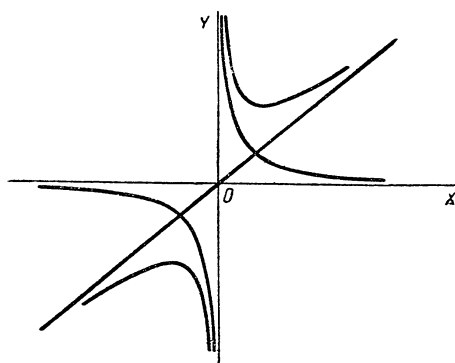


Рис. 138.

Задача 3. Построить график функции $af(x)$, считая известным график функции $f(x)$.

Решение. Пусть $a > 0$ [при $a < 0$ сначала построим график функции $-f(x)$, симметричный графику функции $f(x)$ относительно оси OX]. Для каждого x из области определения функции $f(x)$ соответствующее значение функции $f(x)$ умножается на одно и то же число a ($a > 0$).

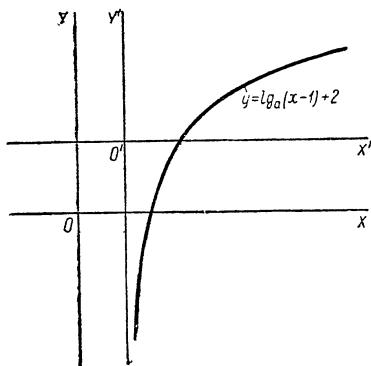


Рис. 139.

Нетрудно видеть, что при $a < 1$ эта операция геометрически представляет собой «сжатие» графика функции $f(x)$ в вертикальном направлении, а при $a > 1$ — «растяжение» графика функции $f(x)$ в вертикальном направлении.

Итак, для построения графика функции $af(x)$ ($a > 0$) нужно график функции $f(x)$ деформировать в a раз вдоль оси OY , т. е. «сжать» при $a < 1$, «растянуть» при $a > 1$. Или, если оставить график функции $f(x)$ без изменения, то нужно изменить в a раз масштаб по оси OY , а именно, уменьшить его при $a > 1$ и увеличить при $a < 1$.

Задача 4. Зная график функции $f(x)$, построить график функции $f(ax)$.
Решение. Пусть $a > 0$ [при $a < 0$ сначала построим график функции $f(-x)$, симметричный графику функции $f(x)$ относительно оси OY].

Функция $f(ax)$ в точке $\frac{x}{a}$ имеет такое же значение, как и функция $f(x)$ в точке x . Учитывая это, получим, что для построения графика функции $f(ax)$ следует график функции $f(x)$ деформировать в a раз вдоль оси OX , т. е. «сжать» при $a > 1$ и «растянуть» при $a < 1$. Или, что то же, не изменяя графика функции $f(x)$, изменить в a раз масштаб по оси OX : уменьшить при $a < 1$ и увеличить при $a > 1$.

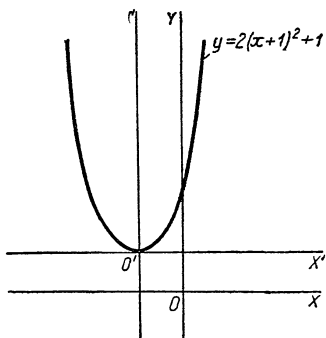


Рис. 140.

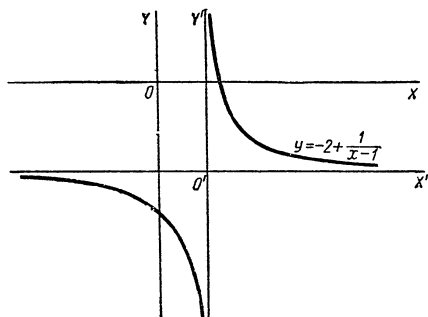


Рис. 141.

Пример 1. Построим график функции $y = 2x^2 + 4x + 3$. Для этого запишем у следующим образом:

$$y = 2 \left(x^2 + 2x + \frac{3}{2} \right) = 2 \left[(x+1)^2 + \frac{1}{2} \right] = 2(x+1)^2 + 1$$

и построение будем производить, например, в таком порядке:

$$1) y = x^2, \quad 2) y = 2x^2, \quad 3) y = 2(x+1)^2, \quad 4) y = 2(x+1)^2 + 1$$

(см. рис. 140)¹.

¹ На рис. 140 парабола $y = x^2$ не показана. Кривые 3) и 4) получены смещением осей.

Пример 2. Построим график функции $y = \frac{-2x+3}{x-1}$.

Запишем $y = \frac{-2x+3}{x-1} = -2 + \frac{1}{x-1}$ и будем производить построение (рис. 141), например, в следующем порядке:

$$1) y = \frac{1}{x}, \quad 2) y = \frac{1}{x-1}, \quad 3) y = -2 + \frac{1}{x-1}.$$

Простая гармоника и ее график. На практике часто встречается функция вида:

$$y = A \sin(\omega x + \alpha), \text{ где } A > 0 \text{ и } \omega > 0.$$

Известно, что такой функцией описывается закон гармонического колебательного движения. Эта функция называется *простой гармоникой*, число A

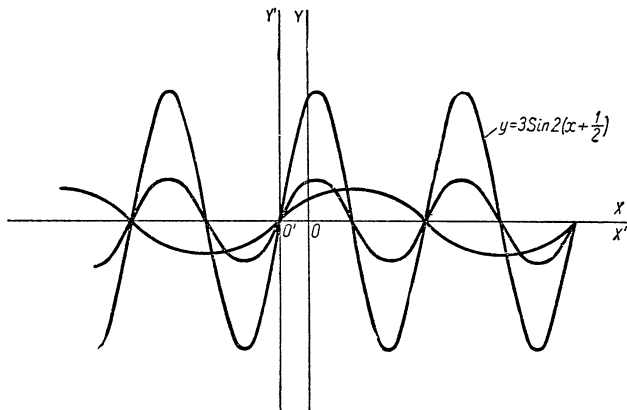


Рис. 142.

называется *амплитудой*, ω — *частотой* и α — *начальной фазой*. График простой гармоникой можно получить из графика синусоиды путем деформаций.

Действительно, запишем $y = A \sin(\omega x + \alpha) = A \sin\left[\omega\left(x + \frac{\alpha}{\omega}\right)\right]$ и последовательно построим графики функций:

$$1) y = \sin x, \quad 2) y = \sin \omega x, \quad 3) y = \sin\left[\omega\left(x + \frac{\alpha}{\omega}\right)\right],$$

$$4) y = A \sin\left[\omega\left(x + \frac{\alpha}{\omega}\right)\right] = A \sin(\omega x + \alpha).$$

Например, построим график функции $y = 3 \sin(2x + 1)$. Записывая y в виде: $y = 3 \sin\left[2\left(x + \frac{1}{2}\right)\right]$, построим последовательно графики функций:

$$1) y = \sin x, \quad 2) y = \sin 2x, \quad 3) y = \sin 2\left[x + \frac{1}{2}\right],$$

$$4) y = 3 \sin\left[2\left(x + \frac{1}{2}\right)\right] \text{ (см. рис. 142).}$$

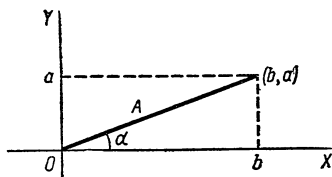


Рис. 143.

Другое аналитическое выражение для простой гармоникой. Так как $A \sin(\omega x + \alpha) =$

$= A \sin \alpha \cos \omega x + A \cos \alpha \sin \omega x$, то, обозначая $A \sin \alpha = a$, $A \cos \alpha = b$, получим: $A \sin(\omega x + \alpha) = a \cos \omega x + b \sin \omega x$. И обратно, если дано выражение $a \cos \omega x + b \sin \omega x$, где a и b — любые действительные числа, то, рассматривая для точки с прямоугольными координатами (b, a) полярные координаты (A, α) (рис. 143), получим $a = A \sin \alpha$, $b = A \cos \alpha$ и, значит,

$$a \cos \omega x + b \sin \omega x = A \sin \alpha \cos \omega x + A \cos \alpha \sin \omega x = A \sin(\omega x + \alpha).$$

Таким образом, простые гармоникой с частотой ω можно записать в виде $a \cos \omega x + b \sin \omega x$, где a и b — действительные числа.

ЧИСЛОВЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

§ 4. Основные понятия

Пусть каждому натуральному числу n поставлено в соответствие число x_n , тогда говорят, что дана *последовательность* $x_1, x_2, \dots, x_n, \dots$, или короче: *последовательность* $\{x_n\}$. *Общий член* последовательности x_n является функцией натурального аргумента n , $x_n = f(n)$. Давая n различные значения: $n = 1, 2, 3, \dots$, получим последовательность значений этой функции:

$$f(1), f(2), \dots, f(n), \dots$$

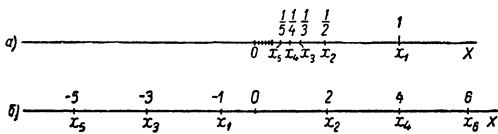
Функцию $f(n)$ можно задать как аналитически, так и любым другим способом.

Например, если $x_n = \frac{1}{n}$, то последовательность запишется в виде:

$$1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$$

Извлекая корень квадратный из двух с недостатком, получаем приближенные значения этого корня с точностью до 1, до $\frac{1}{10}$, до $\frac{1}{100}$ и т. д., т. е. получаем последовательность $x_1 = 1, x_2 = 1,4, x_3 = 1,41, \dots$. Обращая простую дробь, например $\frac{1}{3}$, в десятичную с точностью до одного, двух

и т. д. знаков после запятой, также получаем последовательность:



$$x_1 = 0, \quad x_2 = 0,3 \quad x_3 = 0,33, \dots, \\ x_n = 0,33 \dots 3, \dots,$$

Рис. 144.

при этом в выражении x_n цифра 3 повторяется $(n-1)$ раз.

Таким образом, последовательность задана, если указан способ получения любого ее члена.

Приведем еще примеры последовательностей:

1) $\{x_n\} = \{(-1)^n\}$ или, что то же, $-1, 1, -1, 1, -1, \dots$;

2) $\{x_n\} = \left\{ \sin n \frac{\pi}{2} \right\}$ или, что то же, $1, 0, -1, 0, 1, \dots$;

3) $\{x_n\} = \{(-1)^n n\}$ или, что то же, $-1, 2, -3, 4, -5, \dots$;

4) если $x_n = \begin{cases} 1, & \text{при } n - \text{нечетном,} \\ 0, & \text{при } n - \text{четном,} \end{cases}$

то последовательность имеет вид $1, 0, 1, 0, 1, \dots$;

5) полагая $x_1 = 0, x_2 = 1, x_{n+2} = x_n + x_{n+1}$ ($n = 1, 2, 3, \dots$), т. е. считая каждый член последовательности начиная с третьего равным сумме двух предшествующих ее членов, получим: $0, 1, 1, 2, 3, 5, 8, 13, \dots$

Геометрическое изображение последовательности. Построив на числовой оси точки с абсциссами, равными величинам соответствующих членов последовательности $\{x_n\}$, мы получим последовательность точек, являющуюся геометрическим изображением данной последовательности.

На рис. 144 изображены соответственно последовательности

$$\{x_n\} = \left\{ \frac{1}{n} \right\}, \quad \{x_n\} = \{(-1)^n n\}.$$

Ограниченные и неограниченные последовательности. Последовательность $\{x_n\}$ называется *ограниченной*, если существует число $M > 0$, такое, что для всех n выполняется неравенство

$$|x_n| < M.$$

Геометрически это означает, что все члены последовательности принадлежат интервалу $(-M, M)$ (рис. 145, а).

Последовательность, не являющаяся ограниченной, называется *неограниченной*. Для неограниченной последовательности при любом $M > 0$ найдутся члены последовательности, которые лежат вне интервала $(-M, M)$.

Следует иметь в виду, что в определении ограниченной последовательности ничего не говорится о величине M . Последовательность ограничена, если существует число (обозначим его M), удовлетворяющее условию $|x_n| < M$ при всех n . Число M не обязано быть «наименьшим», удовлетворяющим этому условию.

Последовательность $\{x_n\}$ называется *ограниченной сверху* (говорят также — *справа*), если все ее члены меньше некоторого числа M , т. е. неравенство $x_n < M$ выполняется при любом n , и *ограничена снизу* (слева), если существует число m , такое, что неравенство $x_n > m$ выполняется при всех n .

В случае ограниченности последовательности сверху все ее члены принадлежат интервалу $(-\infty, M)$, в случае ограниченности снизу —

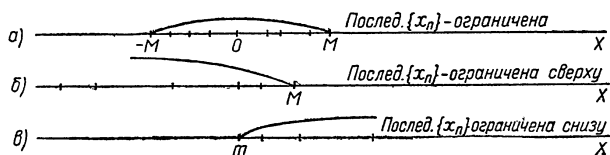


Рис. 145.

интервалу $(m, +\infty)$ (см. рис. 145, б, в). Из этих определений непосредственно следует, что если последовательность ограничена, то она ограничена сверху и снизу, если же последовательность ограничена только с одной стороны, то она может быть неограниченной.

Рассмотрим примеры. Последовательность $\{x_n\} = \{n\}$ ограничена снизу, но она неограниченная, последовательность $\{x_n\} = \{-n\}$ ограничена сверху, но не является ограниченной, и, наконец, последовательность $\{x_n\} = \{(-1)^n n\}$ не ограничена ни сверху, ни снизу.

Последовательность $\{x_n\} = \left\{\frac{1}{n}\right\}$ — ограниченная, так как $|x_n| < 2$ при любом n .

Последовательность $\{x_n\} = \{q^n\}$ ($q > 1$) — неограниченная, так как, положив $q = 1 + \theta$ ($\theta > 0$), получим:

$$q^n = (1 + \theta)^n > 1 + n\theta$$

и, следовательно, при достаточно больших n ее члены будут больше любого наперед заданного числа. Но эта последовательность ограничена снизу, так как неравенство $q^n > 1$ справедливо при любом n .

§ 5. Предел последовательности.

Число a называется *пределом последовательности* $\{x_n\}$, если для любого положительного числа ϵ найдется такой номер N , что для всех $n > N$ выполняется неравенство $|a - x_n| < \epsilon$. Следует обратить особое внимание на то, что ϵ может быть любым положительным числом, в частности, и сколь угодно малым. Неравенство $|a - x_n| < \epsilon$ можно записать в виде $|x_n - a| < \epsilon$ или $-\epsilon < x_n - a < \epsilon$.

Из определения предела следует, что если последовательность $\{x_n\}$ имеет пределом число a , то члены последовательности при всех $n > N^1$ удовлетворяют неравенствам $a - \epsilon < x_n < a + \epsilon$. Интервал $(a - \epsilon, a + \epsilon)$

¹⁾ При надлежащем выборе N .

называют « ϵ -окрестностью» точки a . Учитывая это, определение предела последовательности можно геометрически сформулировать следующим образом: число a является пределом последовательности $\{x_n\}$, если, какова бы ни была ϵ -окрестность точки a , начиная с некоторого номера все точки x_n попадут в эту окрестность, т. е. вне интервала $(a - \epsilon, a + \epsilon)$ останется лишь конечное число членов последовательности (рис. 146). Номер N , о котором говорится в определении предела последовательности, зависит от величины ϵ . Тот факт, что последовательность $\{x_n\}$ имеет предел, равный a , записывают следующим образом: $\lim x_n = a$ ($\lim$ есть сокращение латинского слова *limes*, которое означает «предел»),

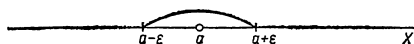


Рис. 146.

$\lim_{n \rightarrow \infty} x_n = a$ или $x_n \rightarrow a$. В этом случае говорят, что последовательность $\{x_n\}$ *сходится* (сходится к a), или имеет конечный предел.

Из определения предела последовательности непосредственно следует, что при отбрасывании, добавлении или изменении конечного числа членов данной последовательности получаются новые последовательности, обладающие следующим свойством: если сходится одна из них, то сходится и другая и их пределы равны.

Это свойство короче можно сформулировать так: две последовательности, отличающиеся между собой на конечное число членов, ведут себя одинаково относительно сходимости.

Рассмотрим примеры.

1) Последовательность $\{x_n\} = \left\{ \frac{(-1)^n}{n} \right\}$ (рис. 147). Докажем, что

$$\lim \frac{(-1)^n}{n} = 0.$$

Действительно, возьмем произвольное положительное число ϵ и определим соответствующий номер N так, чтобы при $n > N$ выполнялось неравенство $\left| 0 - \frac{(-1)^n}{n} \right| < \epsilon$, т. е. $\frac{1}{n} < \epsilon$. Это неравенство выполняется при $n > \frac{1}{\epsilon}$, поэтому за N можно взять целую часть числа $\frac{1}{\epsilon}$. Тогда при всяком $n > N$ выполняется неравенство

$$\left| 0 - \frac{(-1)^n}{n} \right| < \epsilon,$$

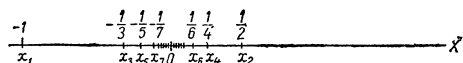


Рис. 147.

а это и означает, что предел последовательности $\{x_n\} = \left\{ \frac{(-1)^n}{n} \right\}$ равен нулю.

Положив, например, $\epsilon = 0,001$, определим $N = 1000$, если $\epsilon = 0,000001$, то $N = 1\,000\,000$.

2) $x_n = q^n \rightarrow 0$ при условии, что $|q| < 1$.

Для доказательства возьмем любое число $\epsilon > 0$ и рассмотрим неравенство

$$\begin{aligned} |0 - q^n| &< \epsilon, \\ |q|^n &< \epsilon. \end{aligned}$$

Чтобы не иметь дела с отрицательными значениями логарифма (так как $|q| < 1$), заменим это неравенство ему равносильным

$$\left(\frac{1}{|q|} \right)^n > \frac{1}{\epsilon}.$$

Тогда, после логарифмирования, получим

$$n \lg \frac{1}{|q|} > \lg \frac{1}{\epsilon},$$

откуда

$$n > \frac{\lg \frac{1}{\epsilon}}{\lg \frac{1}{|q|}}.$$

Возьмем

$$N = E \left(\frac{\lg \frac{1}{\epsilon}}{\lg \frac{1}{|q|}} \right),$$

где $E \left(\frac{\lg \frac{1}{\epsilon}}{\lg \frac{1}{|q|}} \right)$ — целая часть числа $\frac{\lg \frac{1}{\epsilon}}{\lg \frac{1}{|q|}}$. Тогда при $n > N$ выполняется неравенство $|0 - q^n| < \epsilon$, т. е. $q^n \rightarrow 0$.

$$3) \ x_n = \begin{cases} 1 + \frac{1}{n}, & \text{если } n \text{ нечетно,} \\ 2 - \frac{1}{n}, & \text{если } n \text{ четно.} \end{cases}$$

Последовательность $2, \frac{3}{2}, \frac{4}{3}, \frac{7}{4}, \dots$ не имеет предела (см. рис. 148).

Действительно, при n нечетных члены последовательности «сгущаются» к точке 1, при n четных члены последовательности «сгущаются» к точке 2. И ни одно из чисел 1 и 2 не является пределом последовательности, так как, какую бы достаточно малую окрестность точки 1 (или точки 2) ни взять, вне ее содержится бесчисленное множество членов последовательности.

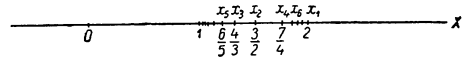


Рис. 148.

4) Последовательности $\{x_n\} = \{(-1)^n\}$, $\{x_n\} = \left\{ \sin n \frac{\pi}{2} \right\}$ также не имеют предела.

Рассмотренные примеры показывают, что если последовательность сходится, то ее члены могут быть как больше, так и меньше предела. Например, у последовательности $\{x_n\} = \left\{ \frac{1}{2^n} \right\}$, сходящейся к 0 (см. пример 2, при $q = \frac{1}{2}$), все члены последовательности положительны, а у последовательности $\{x_n\} = \left\{ \frac{(-1)^n}{2^n} \right\}$ (пример 2, при $q = -\frac{1}{2}$), сходящейся к 0, члены последовательности принимают как положительные, так и отрицательные значения. И, наконец, бесчисленное множество членов последовательности могут быть равными пределу, как это, например, будет для последовательности $\{x_n\}$, определяемой условием

$$x_n = \begin{cases} 1, & \text{если } n \text{ нечетно,} \\ 1 - \frac{1}{n}, & \text{если } n \text{ четно.} \end{cases}$$

Здесь $\lim x_n = 1$.

Из двух членов сходящейся последовательности член с большим номером не всегда будет «ближе» к пределу этой последовательности.

Рассмотрим теперь следующие теоремы.

Теорема 1. Последовательность может иметь не более одного предела.

Пусть $x_n \rightarrow a$, $x_n \rightarrow b$ ($a \neq b$). Возьмем число $\epsilon > 0$. Из того, что $x_n \rightarrow a$, следует, что $|a - x_n| < \frac{\epsilon}{2}$ при $n > N_1$. Из того, что $x_n \rightarrow b$, следует, что $|b - x_n| < \frac{\epsilon}{2}$ при $n > N_2$.

Тогда для $n > N$, где N большее из чисел N_1 и N_2 , выполняются оба неравенства. И, следовательно,

$$|a - b| = |(a - x_n) + (x_n - b)| \leq |a - x_n| + |x_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad |a - b| < \varepsilon.$$

Разность между числами a и b по абсолютной величине меньше любого положительного числа, следовательно, $a = b$.

Теорема 2. Если $x_n \rightarrow a$, то $|x_n| \rightarrow |a|$.

Действительно, для любого $\varepsilon > 0$ неравенство $|a - x_n| < \varepsilon$ будет выполняться при $n > N(\varepsilon)$. Но по свойству абсолютных величин можно записать, что

$$||a| - |x_n|| \leq |a - x_n|,$$

и, следовательно, при $n > N$ выполняется неравенство

$$||a| - |x_n|| < \varepsilon, \text{ т. е. } |x_n| \rightarrow |a|.$$

Теорема 3. Если $x_n \rightarrow a$, то последовательность $\{x_n\}$ ограничена.

Доказательство. Положим $\varepsilon = 1$, тогда найдется номер N такой, что для $n > N$ $|a - a_n| < 1$. Следовательно, $|a_n| = |(a_n - a) + a| \leq |a_n - a| + |a| < 1 + |a|$. Откуда $|a_n| < |a| + 1$ для всех $n > N$. Обозначая через M наибольшее из чисел $|a_1|, |a_2|, \dots, |a_N|, |a| + 1$, получим для всех n : $|a_n| \leq M$, откуда следует, что последовательность $\{x_n\}$ ограничена¹⁾.

Заметим, что обратная теорема неверна. Из ограниченности последовательности не следует ее сходимости, как это видно из рассмотренных ниже примеров:

$$1) \text{ Последовательность } \{x_n\}, x_n = \begin{cases} 1 + \frac{1}{n}, & \text{если } n \text{ четно,} \\ 2 - \frac{1}{n}, & \text{если } n \text{ нечетно,} \end{cases} \text{ не имеет}$$

предела, хотя $|x_n| \leq 2$ при всех n .

2) Ограниченная последовательность $\{x_n\} = \left\{\sin n \frac{\pi}{2}\right\}$, принимающая значения $1, 0, -1, 0, 1, 0, -1 \dots$, также не имеет предела.

Таким образом, ограниченность последовательности является необходимым, но не достаточным условием сходимости последовательности. Последовательность $\{x_n\}$, не имеющая конечного предела, называется *расходящейся*. Следовательно, только что рассмотренные последовательности — расходящиеся. Последовательности $\{(-1)^n\}$, $\{n^2\}$, $\left\{n + \frac{(-1)^n}{n}\right\}$ — также расходящиеся.

§ 6. Бесконечный предел последовательности

Бесконечно большие. Последовательность $\{x_n\}$ называется *бесконечно большой*, если для любого числа $M > 0$ найдется номер N , такой, что для всех $n > N$ выполняется неравенство $|x_n| > M$. Геометрически это означает, что какое бы мы число $M > 0$ ни взяли, все члены последовательности $\{x_n\}$, кроме конечного числа их, лежат вне отрезка $[-M, M]$. Номер N , начиная с которого выполняется неравенство $|x_n| > M$, вообще говоря, зависит от M .

Примерами бесконечно больших последовательностей являются последовательности $\{n\}$, $\{-n\}$, $\{(-1)^{n+1}n\}$. При этом члены первой последовательности принимают только положительные значения, члены второй

¹⁾ Если для всех n $|a_n| \leq M$, то существует число M_1 , равное, например, $M + 1$, для которого выполняется строгое неравенство $|a_n| < M_1$ (см. определение ограниченной последовательности).

последовательности — только отрицательные значения. Члены же третьей последовательности по модулю неограниченно растут, принимая то положительные, то отрицательные значения.

Тот факт, что последовательность $\{x_n\}$ бесконечно большая, записывают следующим образом: $x_n \rightarrow \infty$ (x_n стремится к бесконечности), или $\lim_{n \rightarrow \infty} x_n = \infty$ (предел x_n равен бесконечности), или $\lim_{n \rightarrow \infty} x_n = \infty$.

Если условиться под *окрестностью символа* ∞ понимать совокупность значений x , удовлетворяющих неравенству $|x| > M$, где M — некоторое положительное число, то получим следующее определение бесконечно большой последовательности: последовательность называется бесконечно большой, если в любой окрестности символа ∞ лежат

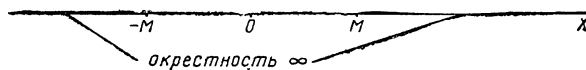


Рис. 149.

все ее члены, кроме, быть может, конечного числа (рис. 149). Всякая бесконечно большая последовательность — неограниченная последовательность.

Иногда бывает важно выделить частные случаи, когда бесконечно большая последовательность $\{x_n\}$ при всех n или при всех n , кроме конечного числа, сохраняет знак. Если последовательность $\{x_n\}$ — бесконечно большая и ее члены, кроме, быть может, конечного числа, положительны, то говорят, что последовательность $\{x_n\}$ стремится к $+\infty$, и пишут $x_n \rightarrow +\infty$, или $\lim_{n \rightarrow \infty} x_n = +\infty$, или $\lim_{n \rightarrow \infty} x_n = +\infty$. Если же члены бесконечно большой последовательности $\{x_n\}$, кроме, может быть, конечного числа их, отрицательны, то говорят, что последовательность $\{x_n\}$ стремится к $-\infty$, и пишут: $x_n \rightarrow -\infty$, или $\lim_{n \rightarrow \infty} x_n = -\infty$, или $\lim_{n \rightarrow \infty} x_n = -\infty$.

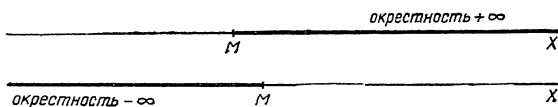


Рис. 150.

Отметим, что записи $x_n \rightarrow +\infty$ и $x_n \rightarrow -\infty$, являющиеся частными случаями (разновидностями) записи $x_n \rightarrow \infty$, вводятся, когда желают подчеркнуть, что члены бесконечно большой последовательности, начиная с некоторого номера, сохраняют определенный знак.

Итак, по определению $\lim_{n \rightarrow \infty} x_n = +\infty$, если для любого числа M найдется номер N , такой, что $x_n > M$ для всех $n > N$ и $\lim_{n \rightarrow \infty} x_n = -\infty$, если для любого числа M найдется номер N , такой, что для всех $n > N$ выполняется неравенство $x_n < M$.

В первом случае, начиная с некоторого номера, все члены последовательности лежат в интервале $(M, +\infty)$, во втором случае — в интервале $(-\infty, M)$. Будем понимать под *окрестностью символа* $+\infty$ совокупность значений x , удовлетворяющих условию $x > M$, а под *окрестностью символа* $-\infty$ совокупность значений x , удовлетворяющих неравенству $x < M$ (рис. 150).

Тогда $\lim_{n \rightarrow \infty} x_n = +\infty$ ($\lim_{n \rightarrow \infty} x_n = -\infty$), если в любой окрестности $+\infty$ ($-\infty$) лежат все члены последовательности, кроме конечного числа их. Число членов, не попавших в выбранную окрестность, зависит от «величины» этой окрестности, т. е. от выбранного числа M .

Рассмотрим еще п р и м е р ы бесконечно больших последовательностей:

$$1) \{x_n\} = \{n^2\}, x_n \rightarrow \infty \text{ и } x_n \rightarrow +\infty,$$

$$2) \{x_n\} = \{(-1)^n n\}, x_n \rightarrow \infty,$$

$$3) \{x_n\} = \left\{ \lg \frac{1}{n} \right\}, x_n \rightarrow \infty \text{ и } x_n \rightarrow -\infty,$$

$$4) \{x_n\} = \left\{ \operatorname{tg} \left(\frac{\pi}{2} + \frac{(-1)^n}{n} \right) \right\}, x_n \rightarrow \infty,$$

$$5) \{x_n\} = \left\{ n + \sin n \frac{\pi}{2} \right\}, x_n \rightarrow \infty \text{ и } x_n \rightarrow +\infty.$$

Теорема о том, что последовательность может иметь не более одного предела, остается верной и после введения бесконечных пределов.

Действительно, двух бесконечных пределов быть не может, а если допустить, что один предел конечен, а другой бесконечен, то придем к противоречию, так как наличие конечного предела последовательности дает ограниченную последовательность и, следовательно, она не может быть бесконечно большой.

Следует иметь в виду, что не всякая неограниченная последовательность является бесконечно большой. Чтобы убедиться в этом, достаточно рассмотреть последовательность $\{x_n\} = \{n^{(-1)^n}\}$, общий член которой можно записать и так:

$$x_n = \begin{cases} n & \text{при } n \text{ четном,} \\ \frac{1}{n} & \text{при } n \text{ нечетном.} \end{cases}$$

Эта последовательность неограниченная и не является бесконечно большой. Она не имеет никакого — ни конечного, ни бесконечного предела. Для неограниченной последовательности логически имеются две возможности: она или имеет бесконечный предел, т. е. является бесконечно большой, или не имеет никакого предела.

Выше была определена *сходящаяся* последовательность, как последовательность, имеющая конечный предел, в противном случае последовательность называлась *расходящейся*. Таким образом, последовательность расходится, если она не имеет предела или ее предел равен бесконечности.

§ 7. Бесконечно малые

Последовательность $\{x_n\}$ называется *бесконечно малой*, или нуль-последовательностью, если $\lim x_n = 0$. Например, последовательности $\{x_n\} = \left\{ \frac{1}{n^2} \right\}$, $\{x_n\} = \left\{ \sin \frac{\pi}{n} \right\}$ — бесконечно малые. Последовательность $\{x_n\} = \{0\}$, все члены которой равны нулю, также бесконечно малая. Общие члены бесконечно малых последовательностей будем обозначать буквами греческого алфавита, например $\alpha_n, \beta_n, \gamma_n, \dots$. Из определения предела последовательности непосредственно получаем, что последовательность $\{a_n\}$ будет бесконечно малой, если для любого положительного числа ϵ найдется число N , такое, что для $n > N$ выполняется неравенство $|\alpha_n| < \epsilon$.

Если $\lim x_n = a$, то, как это следует из определения предела последовательности, разность $x_n - a = \alpha_n$, где $\alpha_n \rightarrow 0$. Обратно, если разность $x_n - a = \alpha_n$, где $\alpha_n \rightarrow 0$, то $\lim x_n = a$. Таким образом, установлена связь между пределами сходящихся последовательностей и бесконечно малыми.

Теорема. Для того чтобы последовательность $\{x_n\}$ сходилась к числу a , необходимо и достаточно, чтобы выполнялось равенство

$$x_n = a + \alpha_n, \text{ где } \alpha_n \rightarrow 0.$$

Связь между бесконечно малыми и бесконечно большими.

Теорема. Если последовательность $\{\alpha_n\}$ бесконечно малая и все ее члены отличны от нуля, то последовательность $\{x_n\} = \left\{\frac{1}{\alpha_n}\right\}$ — бесконечно большая, и обратно, если последовательность $\{x_n\}$, $x_n \neq 0$, бесконечно большая, то последовательность $\{\alpha_n\} = \left\{\frac{1}{x_n}\right\}$ — бесконечно малая.

Действительно, пусть $\alpha_n \rightarrow 0$, тогда для любого числа $M > 0$ найдется номер N , такой, что при $n > N$ выполняется неравенство $|\alpha_n| < \frac{1}{M}$ или $\left|\frac{1}{\alpha_n}\right| > M$, т. е. $x_n = \frac{1}{\alpha_n} \rightarrow \infty$. Пусть теперь $x_n \rightarrow \infty$, тогда для любого числа $\varepsilon > 0$ найдется число N , такое, что неравенство $|x_n| > \frac{1}{\varepsilon}$ выполняется при $n > N$, откуда получим, что $\frac{1}{|x_n|} < \varepsilon$ для всех $n > N$, т. е. $\alpha_n = \frac{1}{x_n} \rightarrow 0$.

Например, последовательность $\{q^n\}$ ($|q| < 1$) — бесконечно малая, последовательность $\{Q^n\}$ ($|Q| > 1$) — бесконечно большая.

Операции над последовательностями. Суммой, разностью, произведением и частным двух последовательностей $\{x_n\}$, $\{y_n\}$ называются соответственно последовательности

$$\{x_n + y_n\}, \{x_n - y_n\}, \{x_n \cdot y_n\}, \left\{\frac{x_n}{y_n}\right\},$$

при делении следует сделать оговорку, что все члены последовательности-делителя отличны от нуля.

Произведение последовательности $\{x_n\}$ на число C можно рассматривать как произведение последовательности $\{x_n\}$ на последовательность $\{y_n\} = \{C\}$, все члены которой равны C , в результате получим последовательность $\{Cx_n\}$.

Теоремы о бесконечно малых.

Теорема 1. Алгебраическая сумма двух бесконечно малых есть бесконечно малая.

Пусть $\alpha_n \rightarrow 0$ и $\beta_n \rightarrow 0$. Докажем, что $\alpha_n \pm \beta_n \rightarrow 0$. Возьмем произвольное число $\varepsilon > 0$, тогда существует номер N_1 , такой, что при $n > N_1$ имеем $|\alpha_n| < \frac{\varepsilon}{2}$, и номер N_2 , такой, что при $n > N_2$ имеем $|\beta_n| < \frac{\varepsilon}{2}$.

Обозначим через N наибольшее из чисел N_1 и N_2 , тогда при $n > N$ выполняются оба неравенства и, следовательно,

$$|\alpha_n \pm \beta_n| \leq |\alpha_n| + |\beta_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Так как для любого числа $\varepsilon > 0$, $|\alpha_n \pm \beta_n| < \varepsilon$ при $n > N$, то последовательность $\{\alpha_n \pm \beta_n\}$ — бесконечно малая, что и требовалось доказать.

С помощью этой теоремы, если применить ее последовательно несколько раз, получаем, что алгебраическая сумма любого фиксированного числа бесконечно малых есть бесконечно малая.

Здесь очень важно помнить, что число слагаемых фиксировано. Теорема перестает быть верной, если число бесконечно малых слагаемых неограниченно растет.

Например, в случае, когда имеется n слагаемых, равных $\frac{1}{n}$ ($\frac{1}{n} \rightarrow 0$), то их сумма

$\underbrace{\frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}}_n$, обозначим ее через x_n , равна 1, и, следовательно, $x_n \rightarrow 1$, если слагаемых, равных $\frac{1}{n}$, будет n^2 , то их сумма x_n равна n , $x_n = n$ и, следовательно, $x_n \rightarrow \infty$.

Теорема 2. Произведение ограниченной последовательности и бесконечно малой последовательности есть бесконечно малая последовательность.

Пусть последовательность $\{x_n\}$ — ограниченная, а $\{\alpha_n\}$ бесконечно малая. Докажем, что $x_n \alpha_n \rightarrow 0$.

В самом деле, существует число $M > 0$, такое, что неравенство $|x_n| < M$ выполняется при всех n . Возьмем любое число $\varepsilon > 0$, ему соответствует число N , такое, что при $n > N$ $|\alpha_n| < \frac{\varepsilon}{M}$. Тогда $|x_n \alpha_n| < M \cdot \frac{\varepsilon}{M} = \varepsilon$, т. е. $|x_n \alpha_n| < \varepsilon$ при $n > N$, а это и означает, что $x_n \alpha_n \rightarrow 0$.

Следствие 1. Произведение сходящейся последовательности и бесконечно малой последовательности есть бесконечно малая последовательность.

Действительно, это утверждение получается сразу же, если учесть свойство ограниченности сходящейся последовательности.

Следствие 2. Произведение двух бесконечно малых последовательностей есть бесконечно малая последовательность.

Произведение любого фиксированного числа бесконечно малых — также бесконечно малая.

Заметим, что об отношении двух бесконечно малых, а также о произведении бесконечно большой на бесконечно малую ничего определенного сказать нельзя.

Рассмотрим примеры.

$$\text{I. 1) } \alpha_n = \frac{1}{n} \rightarrow 0, \quad \beta_n = \frac{1}{n^2} \rightarrow 0, \quad \frac{\alpha_n}{\beta_n} = n \rightarrow \infty, \quad \frac{\beta_n}{\alpha_n} = \frac{1}{n} \rightarrow 0.$$

$$2) \alpha_n = \frac{(-1)^n}{n} \rightarrow 0, \quad \beta_n = \frac{1}{n} \rightarrow 0, \quad \frac{\alpha_n}{\beta_n} = (-1)^n \text{ — предела не существует.}$$

$$3) \alpha_n = \frac{1}{n} \rightarrow 0, \quad \beta_n = \frac{2}{n} \rightarrow 0, \quad \frac{\alpha_n}{\beta_n} = \frac{1}{2} \rightarrow \frac{1}{2}.$$

$$\text{II. 1) } \alpha_n = \frac{1}{n} \rightarrow 0, \quad \beta_n = n \rightarrow \infty, \quad \alpha_n \beta_n = 1 \rightarrow 1.$$

$$2) \alpha_n = \frac{1}{n} \rightarrow 0, \quad \beta_n = n^2 \rightarrow \infty, \quad \alpha_n \beta_n = n \rightarrow \infty.$$

$$3) \alpha_n = \frac{1}{n^2} \rightarrow 0, \quad \beta_n = n \rightarrow \infty, \quad \alpha_n \beta_n = \frac{1}{n} \rightarrow 0.$$

$$4) \alpha_n = \frac{(-1)^n}{n} \rightarrow 0, \quad \beta_n = n \rightarrow \infty, \quad \alpha_n \beta_n = (-1)^n \text{ — не имеет предела.}$$

§ 8. Теоремы о пределах

Теорема 1. Алгебраическая сумма двух сходящихся последовательностей есть сходящаяся последовательность, ее предел равен соответствующей сумме пределов данных последовательностей.

Пусть $x_n \rightarrow a$, $y_n \rightarrow b$. Докажем, что $x_n \pm y_n \rightarrow a \pm b$. Действительно, $x_n = a + \alpha_n$, где $\alpha_n \rightarrow 0$, $y_n = b + \beta_n$, где $\beta_n \rightarrow 0$. Тогда $x_n \pm y_n = (a \pm b) + (\alpha_n \pm \beta_n)$. По доказанной теореме о сумме бесконечно малых последовательностей $\{\alpha_n \pm \beta_n\}$ — бесконечно малая, окончательно получим

$$\lim (x_n \pm y_n) = a \pm b, \text{ т. е. } \\ \lim (x_n \pm y_n) = \lim x_n \pm \lim y_n.$$

Если условия теоремы выполняются, то теорему будем формулировать короче: предел суммы равен сумме пределов¹⁾.

Теорема остается верной и в случае любого фиксированного числа сходящихся последовательностей.

Теорема 2. Произведение двух сходящихся последовательностей есть сходящаяся последовательность, ее предел равен произведению пределов данных последовательностей.

Пусть $x_n \rightarrow a$, $y_n \rightarrow b$, докажем, что $x_n y_n \rightarrow ab$. Действительно,

$$\begin{aligned}x_n &= a + \alpha_n, \quad \alpha_n \rightarrow 0 \\y_n &= b + \beta_n, \quad \beta_n \rightarrow 0.\end{aligned}$$

Тогда $x_n y_n = (a + \alpha_n)(b + \beta_n) = ab + (a\beta_n + b\alpha_n + \alpha_n \beta_n)$. На основании свойств бесконечно малых заключаем, что $(a\beta_n + b\alpha_n + \alpha_n \beta_n) \rightarrow 0$, т. е. последовательность $\{a\beta_n + b\alpha_n + \alpha_n \beta_n\}$ — бесконечно малая. Отсюда

$$\lim (x_n y_n) = ab, \text{ т. е. } \lim (x_n y_n) = \lim x_n \lim y_n.$$

Если условия теоремы выполняются, то предел произведения равен произведению пределов²⁾.

Теорема остается справедливой и в случае произведения любого фиксированного числа сходящихся последовательностей.

Следствие 1. Если последовательность $\{x_n\}$ имеет конечный предел, то при любом постоянном C имеем: $\lim Cx_n = C \lim x_n$, т. е. постоянный множитель можно выносить за знак предела.

Чтобы доказать теорему о пределе частного двух сходящихся последовательностей, рассмотрим две леммы.

Лемма 1. Если $\lim x_n = a$, $a \neq 0$, то существует число $K > 0$ и натуральное число N , такие, что при $n > N$ выполняется неравенство $|x_n| > K$. В этом случае говорят, что последовательность $\{x_n\}$ отделима от нуля.

Доказательство. По условию $\lim x_n = a$, тогда как известно (теорема 2, § 5), $\lim |x_n| = |a| > 0$ и, следовательно, для любого числа $\epsilon > 0$, начиная с некоторого номера $N = N(\epsilon)$, т. е. для $n > N$, выполняются неравенства $|a| - \epsilon < x_n < |a| + \epsilon$.

Возьмем ϵ , удовлетворяющее условию $0 < \epsilon < \frac{|a|}{2}$. Тогда, учитывая левое неравенство, получим, что при $n > N$, $|x_n| > \frac{|a|}{2}$. Полагая $K = \frac{|a|}{2}$, получим утверждение леммы.

В частности, из этой леммы следует, что если предел последовательности не равен нулю, то, начиная с некоторого номера, все члены последовательности отличны от нуля.

Лемма 2. Если $\lim x_n = a$, $a \neq 0$, и все члены последовательности отличны от нуля, то последовательность $\left\{\frac{1}{x_n}\right\}$ ограничена.

Действительно, по лемме 1 существует число $K > 0$, такое, что при $n > N$ имеет место неравенство $|x_n| > K$, а следовательно, $\left|\frac{1}{x_n}\right| < \frac{1}{K}$. Если обозначим через M наибольшее из чисел $\left|\frac{1}{x_1}\right|, \left|\frac{1}{x_2}\right|, \dots, \left|\frac{1}{x_N}\right|, \frac{1}{K}$, то $\left|\frac{1}{x_n}\right| \leq M$ для всех n ³⁾.

¹⁾ Заметим, что из существования предела суммы двух последовательностей не следует существование предела каждой последовательности, что легко видеть на примере последовательностей

$$\{x_n\} = \{(-1)^n\}, \quad \{y_n\} = \left\{\frac{1}{n} + (-1)^{n+1}\right\}.$$

Здесь $x_n + y_n = \frac{1}{n} \rightarrow 0$, т. е. сумма $\{x_n + y_n\}$ имеет предел, а последовательности $\{x_n\}, \{y_n\}$ не имеют пределов.

²⁾ Для произведения можно сделать замечание, аналогичное замечанию в сноске¹⁾.

³⁾ См. сноску на стр. 162.

Теорема 3. Если последовательности $\{x_n\}$ и $\{y_n\}$ сходятся, $\lim x_n = a$, $\lim y_n = b$, причем $y_n \neq 0$, $n = 1, 2, \dots$ и $b \neq 0$, то последовательность $\left\{\frac{x_n}{y_n}\right\}$ сходится и ее предел равен отношению пределов последовательностей $\{x_n\}$ и $\{y_n\}$.

Пусть $\lim x_n = a$, $\lim y_n = b$. При этом $y_n \neq 0$, $n = 1, 2, \dots$, и $b \neq 0$. Докажем, что $\frac{x_n}{y_n} \rightarrow \frac{a}{b}$.

Действительно, $x_n = a + \alpha_n$, $\alpha_n \rightarrow 0$, $y_n = b + \beta_n$, $\beta_n \rightarrow 0$. Тогда $\frac{x_n}{y_n} - \frac{a}{b} = \frac{bx_n - ay_n}{by_n} = \frac{b(a + \alpha_n) - a(b + \beta_n)}{by_n} = \frac{b\alpha_n - a\beta_n}{by_n}$. Так как последовательность $\left\{\frac{1}{y_n}\right\}$ ограничена (лемма 2), а $\frac{b\alpha_n - a\beta_n}{b} \rightarrow 0$ (см. теоремы 1 и 2, § 7), то последовательность $\left\{\frac{b\alpha_n - a\beta_n}{by_n}\right\}$ — бесконечно малая (по теореме 2). Окончательно получаем, что

$$\frac{x_n}{y_n} \rightarrow \frac{a}{b} \quad \text{или} \quad \lim \frac{x_n}{y_n} = \frac{\lim x_n}{\lim y_n}.$$

О неопределенностях. При рассмотрении теорем о пределах мы не рассматривали последовательности, имеющие бесконечные пределы, а также случай, когда при отыскании предела частного предел последовательности, стоящей в знаменателе, равен нулю. Если обе последовательности $\{x_n\}$ и $\{y_n\}$ — бесконечно большие одного знака или одна из них бесконечно большая, а другая ограниченная, то их сумма $\{x_n + y_n\}$ — бесконечно большая последовательность. Этот результат непосредственно вытекает из определения бесконечно большой. Если последовательности $\{x_n\}$ и $\{y_n\}$ — бесконечно большие разных знаков, то про последовательность $\{x_n + y_n\}$ ничего определенного сказать нельзя. Это следует из рассмотренных ниже примеров.

- Примеры: 1) $x_n = n^2 \rightarrow +\infty$, $y_n = -n^2 \rightarrow -\infty$, $x_n + y_n = 0 \rightarrow 0$.
 2) $x_n = n^2 \rightarrow +\infty$, $y_n = -n^2 + 5 \rightarrow -\infty$, $x_n + y_n = 5 \rightarrow 5$.
 3) $x_n = n^2 \rightarrow +\infty$, $y_n = -n^2 + \frac{1}{n} \rightarrow -\infty$, $x_n + y_n = \frac{1}{n} \rightarrow 0$.
 4) $x_n = n^2 \rightarrow +\infty$, $y_n = -n^2 + (-1)^n \rightarrow -\infty$, $x_n + y_n = (-1)^n$ — предела не имеет.

Приведенные примеры показывают, что сумма двух бесконечно больших разных знаков может стремиться к любому числу, к бесконечности, а может вообще не иметь предела. В этом случае говорят, что имеет место *неопределенность*. Ее называют «неопределенностью вида $\infty - \infty$ ».

«Раскрыть» эту неопределенность — значит в каждом конкретном случае, в зависимости от вида последовательностей $\{x_n\}$, $\{y_n\}$, решить вопрос о поведении последовательности $\{x_n + y_n\}$.

При рассмотрении последовательности $\{x_n \cdot y_n\}$ следует обратить внимание на следующую **теорему**: если последовательность $\{x_n\}$ отделима от нуля, а последовательность $\{y_n\}$ — бесконечно большая, то последовательность $\{x_n y_n\}$ — бесконечно большая. Доказательство этой теоремы рекомендуется провести самостоятельно.

В частности, произведение последовательности, сходящейся к a ($a \neq 0$), и бесконечно большой последовательности есть бесконечно большая последовательность. В случае произведения бесконечно малой на бесконечно большую мы получим так называемую «неопределенность вида $0 \cdot \infty$ », смысл которой «раскрывается» при рассмотрении поведения данных последовательностей. Примеры такого вида неопределенности были рассмотрены раньше. Если же обе последовательности бесконечно большие, то их произведение — также бесконечно большая последовательность.

При делении последовательностей могут встретиться «неопределенности» вида $\frac{0}{0}$ и $\frac{\infty}{\infty}$. Выше было показано, что отношение двух бесконечно малых является неопределенностью. Отношение двух бесконечно больших — также «неопределенность», как это будет видно на примерах 1)–3). В этом также легко убедиться, если рассматривать отношение двух бесконечно больших как произведение бесконечно малой и бесконечно большой, которое, как известно, дает «неопределенность» вида $0 \cdot \infty$.

- 1) $x_n = n \rightarrow \infty$, $\frac{x_n}{y_n} = \frac{1}{n} \rightarrow 0$, $\frac{y_n}{x_n} = n \rightarrow \infty$.
- 2) $x_n = n \rightarrow \infty$, $\frac{x_n}{y_n} = \frac{1}{2} \rightarrow \frac{1}{2}$.
- 3) $x_n = n \rightarrow \infty$, $\frac{x_n}{y_n} = (-1)^n$ — предела не имеет.

Рассмотрим еще следующие примеры.

Пример 1. Найти $\lim (\sqrt{n} - \sqrt{n+1})$. Здесь нужно раскрыть неопределенность вида $\infty - \infty$.

$$\lim_{n \rightarrow \infty} (\sqrt{n} - \sqrt{n+1}) = \lim_{n \rightarrow \infty} \frac{n - (n+1)}{\sqrt{n} + \sqrt{n+1}} = 0.$$

Пример 2. Найти $\lim_{n \rightarrow \infty} \frac{n^4 + 5n - 3}{n^3 + 7n + 6}$. В этом примере имеем неопределенность вида $\frac{\infty}{\infty}$, «раскроем» ее.

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{n^4 + 5n - 3}{n^3 + 7n + 6} &= \lim_{n \rightarrow \infty} \frac{n^4 \left(1 + \frac{5}{n^3} - \frac{3}{n^4}\right)}{n^3 \left(1 + \frac{7}{n^2} + \frac{6}{n^3}\right)} = \\ &= \lim_{n \rightarrow \infty} n \cdot \frac{1 + \frac{5}{n^3} - \frac{3}{n^4}}{1 + \frac{7}{n^2} + \frac{6}{n^3}}. \end{aligned}$$

Так как

$$\lim_{n \rightarrow \infty} \frac{1 + \frac{5}{n^3} - \frac{3}{n^4}}{1 + \frac{7}{n^2} + \frac{6}{n^3}} = 1,$$

то данная последовательность — бесконечно большая, т. е.

$$\lim_{n \rightarrow \infty} \frac{n^4 + 5n - 3}{n^3 + 7n + 6} = \infty.$$

§ 9. О переходе к пределу в неравенствах

Теорема 1. Если последовательность $\{x_n\}$ сходится и $x_n \geq 0$ при любом n , то $\lim x_n \geq 0$.

Доказательство. Допустим, что $\lim x_n = a < 0$. Тогда для числа $\varepsilon > 0$ найдется номер N , такой, что при $n > N$ будет выполняться неравенство $|a - x_n| < \varepsilon$, или, что то же, $a - \varepsilon < x_n < a + \varepsilon$.

Возьмем ε , удовлетворяющее условию $0 < \varepsilon < \frac{|a|}{2}$. Тогда на основании правого неравенства получим, что $x_n < \frac{a}{2} < 0$ при $n > N$. А это противоречит условию теоремы.

В частности, если последовательность $\{x_n\}$ сходится и при всех n выполняется неравенство $x_n > 0$, то $\lim x_n \geq 0$. Следует иметь в виду, что

предел последовательности может быть равен нулю, и в случае, когда все члены последовательности положительны, как, например, для последовательности $\left\{\frac{1}{n}\right\}$, предел которой равен нулю.

Следствие. Если члены сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$ при всех n удовлетворяют условиям $x_n \leq y_n$ (в частности, если $x_n < y_n$), то $\lim x_n \leq \lim y_n$.

Для доказательства достаточно рассмотреть последовательность $\{y_n - x_n\}$, члены которой неотрицательны (положительны). Тогда по теореме 1

$$\lim (y_n - x_n) = \lim y_n - \lim x_n \geq 0,$$

следовательно, $\lim y_n \geq \lim x_n$.

Теорема 2. Если последовательности $\{x_n\}$ и $\{z_n\}$ сходятся и имеют равные пределы, $\lim x_n = \lim z_n = a$, и если при любом n

$$x_n \leq y_n \leq z_n,$$

то последовательность $\{y_n\}$ сходится и ее предел равен a .

Доказательство. Для любого числа $\varepsilon > 0$ найдется такой номер N_1 , что $|a - x_n| < \varepsilon$ при $n > N_1$, и такой номер N_2 , что $|a - z_n| < \varepsilon$ при $n > N_2$. При $n > N$, где N наибольшее из чисел N_1 и N_2 , выполняются оба эти неравенства, откуда следует, что члены последовательностей $\{x_n\}$ и $\{z_n\}$ с номерами $n > N$ лежат между $a - \varepsilon$ и $a + \varepsilon$. Так как y_n при всех n лежит между x_n и z_n , то все члены последовательности $\{y_n\}$ с номерами $n > N$ лежат между $a - \varepsilon$ и $a + \varepsilon$, или, что то же,

$$a - \varepsilon < y_n < a + \varepsilon.$$

Следовательно, $|a - y_n| < \varepsilon$ при $n > N$, т. е. $\lim y_n = a$.

Для равенства пределов двух сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$ необходимо и достаточно, чтобы их разность была бесконечно малой.

Действительно, обозначим: $x_n - y_n = z_n$. Тогда $\lim x_n - \lim y_n = \lim z_n$. И, следовательно, $\lim x_n = \lim y_n$ лишь при условии: $\lim z_n = 0$, т. е. $\{z_n\} = \{x_n - y_n\}$ — бесконечно малая последовательность.

Если сходящаяся последовательность $\{x_n\}$ имеет предел, равный a , а последовательность $\{y_n\}$ такова, что разность $\{y_n - x_n\}$ является бесконечно малой последовательностью, то $y_n \rightarrow a$. В самом деле, $x_n = a + \alpha_n$, $y_n = x_n + \beta_n$, откуда $y_n = a + (\alpha_n + \beta_n)$. Учитывая, что $\alpha_n + \beta_n \rightarrow 0$, получим $y_n \rightarrow a$.

§ 10. Монотонные последовательности

Последовательность $\{x_n\}$ называется *возрастающей*, если $x_1 < x_2 < \dots < x_n < x_{n+1} < \dots$; *неубывающей*, если $x_1 \leq x_2 \leq \dots \leq x_n \leq x_{n+1} \leq \dots$; *убывающей*, если $x_1 > x_2 > \dots > x_n > x_{n+1} > \dots$; *невозрастающей*, если $x_1 \geq x_2 \geq \dots \geq x_n \geq x_{n+1} \geq \dots$.

Все такие последовательности называются *монотонными*.

Из определения монотонных последовательностей непосредственно следует: если монотонно возрастающая последовательность ограничена сверху, то она ограничена, если монотонно убывающая последовательность ограничена снизу, то она ограничена.

Как известно, в случае немонотонной последовательности из ограниченности последовательности с одной стороны не следует ограниченность этой последовательности.

Теорема. Монотонная ограниченная последовательность сходится.

Доказательство этой теоремы будем вести на основании следующего свойства действительных чисел¹⁾: если последовательность отрезков

$$[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n], \dots$$

¹⁾ В связи с тем что теория действительных чисел нами не рассматривается, это свойство действительных чисел принимается без доказательства.

такова, что каждый следующий принадлежит предыдущему

$$a_n \leq a_{n+1} < b_{n+1} \leq b_n \quad (n = 1, 2, \dots)$$

(в этом случае говорят, что имеется *последовательность вложенных отрезков*), то существует по крайней мере одна точка ξ , общая всем отрезкам, $a_n \leq \xi \leq b_n$ ($n = 1, 2, \dots$).

Если же последовательность вложенных отрезков $[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n], \dots$ удовлетворяет условию $\lim (b_n - a_n) = 0$, т. е. длины отрезков стремятся к нулю, то такая последовательность отрезков называется *последовательностью вложенных стягивающихся отрезков*.

Лемма. Для последовательности вложенных стягивающихся отрезков существует единственная точка, принадлежащая всем отрезкам.

Эта лемма в дальнейшем будет неоднократно использоваться.

Доказательство. По упомянутому свойству действительных чисел для последовательности вложенных отрезков существует по крайней мере одна точка ξ , общая всем отрезкам.

Допустив, что существует еще точка ξ_1 ($\xi_1 \neq \xi$), принадлежащая всем отрезкам, придем к противоречию, так как при всех n

$$b_n - a_n \geq |\xi_1 - \xi|$$

и, следовательно, $\lim (b_n - a_n) \geq |\xi_1 - \xi| \neq 0$, что противоречит условию.

Заметим, что лемма будет неверна, если вместо отрезков рассматривать интервалы. Например, для последовательности интервалов

$$(0, 1), \left(0, \frac{1}{2}\right), \left(0, \frac{1}{2^2}\right), \dots, \left(0, \frac{1}{2^n}\right), \dots,$$

не существует точки, общей всем интервалам.

Доказательство теоремы. Пусть неубывающая последовательность $x_1, x_2, \dots, x_n, \dots$ ограничена сверху. Докажем, что она сходится и все ее члены не превосходят ее предела.

Так как последовательность ограничена сверху, то существует число M , такое, что $x_n < M$ при всех n . Обозначим отрезок $[x_1, M]$ через $[a_1, b_1]$, а его длину — через d . Разделим отрезок $[a_1, b_1]$ пополам, тогда в силу монотонности последовательности $\{x_n\}$ в одном из полученных отрезков содержатся все члены последовательности, кроме, возможно, конечного числа их. Рассмотрим этот отрезок и обозначим его через $[a_2, b_2]$. Разделим теперь отрезок $[a_2, b_2]$ пополам, тогда в одном из полученных отрезков содержатся все члены последовательности, кроме, возможно, конечного числа их, этот отрезок обозначим через $[a_3, b_3]$, и т. д.

Продолжая неограниченно этот процесс деления, получим последовательность $[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n], \dots$ вложенных стягивающихся отрезков (так как длина отрезка $[a_n, b_n]$, равная $\frac{d}{2^{n-1}}$, стремится к нулю при неограниченном возрастании n). В каждом таком отрезке содержатся все члены последовательности, кроме, быть может, конечного числа их. На основании леммы существует единственная точка (обозначим ее через a), принадлежащая всем отрезкам.

Докажем, что $x_n \rightarrow a$. Возьмем произвольное число $\varepsilon > 0$. Тогда в интервал $(a - \varepsilon, a + \varepsilon)$ попадет отрезок $[a_n, b_n]$ достаточно большого номера.

Вне этого отрезка содержится лишь конечное число членов последовательности $\{x_n\}$. Следовательно, для всех членов последовательности, начиная с некоторого номера, выполняется неравенство $|x_n - a| < \varepsilon$, т. е. $x_n \rightarrow a$.

Докажем теперь, что при всех n выполнено неравенство $x_n \leq a$.

Если допустить, что при каком-то значении n_0 $x_{n_0} > a$, то для всех номеров m , больших n_0 , будет $x_m \geq x_{n_0}$ и, следовательно, $x_m - a \geq x_{n_0} - a$, т. е. разность $x_m - a$ не стремится к нулю, что противоречит определе-

нию предела. Итак, для любого n имеем $x_n \leq a$, что и требовалось доказать.

Пусть теперь последовательность $\{x_n\}$ невозрастающая и ограничена снизу. Докажем, что она сходится и при этом все члены последовательности не меньше ее предела.

Действительно, рассмотрим последовательность $\{-x_n\}$. Она неубывающая, ограниченная сверху и, следовательно, сходящаяся.

Обозначим ее предел через $-a$, $\lim(-x_n) = -a$.

Так как для любого $\varepsilon > 0$ существует число N , такое, что при $n \geq N$ выполняется неравенство $|-a - (-x_n)| < \varepsilon$ или, что то же, $|a - x_n| < \varepsilon$, то $\lim x_n = a$. Для неубывающей последовательности $\{-x_n\}$ при любом n справедливо соотношение $-x_n \leq -a$, следовательно, для данной последовательности $x_n \geq a$, что и требовалось доказать. Таким образом, доказана теорема о сходимости любой монотонной ограниченной последовательности¹⁾.

Отметим еще, что всякая монотонная последовательность имеет предел, конечный или бесконечный. В самом деле, если эта последовательность ограничена, то по доказанной теореме она сходится, т. е. имеет конечный предел, если же последовательность не ограниченная, то, очевидно, ее предел равен ∞ ($+\infty$ для монотонно возрастающей и $-\infty$ для монотонно убывающей последовательности)²⁾.

Доказанная теорема о монотонных последовательностях практически во многих случаях позволяет установить существование предела последовательности.

С помощью этой теоремы, например, доказывается существование предела последовательности периметров (площадей) вписанных в окружность правильных n -угольников при неограниченном удвоении числа их сторон и периметров (площадей) описанных многоугольников.

§ 11. Число e .

Рассмотрим последовательность с общим членом $x_n = \left(1 + \frac{1}{n}\right)^n$:

$$(1+1)^1, \left(1+\frac{1}{2}\right)^2, \dots, \left(1+\frac{1}{n}\right)^n, \dots$$

и докажем, что она имеет конечный предел. Для доказательства существования конечного предела этой последовательности достаточно показать, что последовательность $\{x_n\}$ — возрастающая и ограниченная сверху.

Доказательство. Пользуясь формулой бинома Ньютона, запишем:

$$\begin{aligned} x_n &= \left(1 + \frac{1}{n}\right)^n = 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{1 \cdot 2} \frac{1}{n^2} + \\ &+ \dots + \frac{n(n-1)(n-2) \dots [n-(k-1)]}{1 \cdot 2 \cdot 3 \dots k} \frac{1}{n^k} + \\ &+ \dots + \frac{n(n-1)(n-2) \dots [n-(n-1)]}{1 \cdot 2 \cdot 3 \dots n} \frac{1}{n^n} \end{aligned}$$

или, что то же,

$$\begin{aligned} x_n &= 1 + 1 + \frac{1}{2!} \left(1 - \frac{1}{n}\right) + \dots + \frac{1}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) + \\ &+ \dots + \frac{1}{n!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{n-1}{n}\right). \end{aligned} \quad (1)$$

¹⁾ Мы считали известной лемму о вложенных стягивающихся отрезках, доказательство которой приводится в теории действительных чисел. Можно, приняв доказанную теорему за исходную, доказать эту лемму.

²⁾ Заметим, что для неограниченных немонотонных последовательностей возможны два случая: или последовательность имеет бесконечный предел, или последовательность не имеет предела (см. § 6).

Покажем, что последовательность $\{x_n\}$ — возрастающая. Действительно, запишем выражение x_{n+1} и сравним его с выражением x_n :

$$\begin{aligned} x_{n+1} = & 1 + 1 + \frac{1}{2!} \left(1 - \frac{1}{n+1}\right) + \dots + \frac{1}{k!} \left(1 - \frac{1}{n+1}\right) \left(1 - \frac{2}{n+1}\right) \dots \\ & \dots \left(1 - \frac{k-1}{n+1}\right) + \dots + \frac{1}{n!} \left(1 - \frac{1}{n+1}\right) \left(1 - \frac{2}{n+1}\right) \dots \\ & \dots \left(1 - \frac{n-1}{n+1}\right) + \frac{1}{(n+1)!} \left(1 - \frac{1}{n+1}\right) \dots \left(1 - \frac{n}{n+1}\right). \end{aligned}$$

Каждое слагаемое в выражении x_{n+1} больше соответствующего слагаемого в выражении x_n , и, кроме того, у x_{n+1} добавляется еще одно положительное слагаемое. Таким образом, последовательность $\{x_n\}$ возрастающая.

Докажем теперь, что при любом n ее члены не превосходят трех:

$x_n < 3$.

В самом деле,

$$\begin{aligned} x_n < 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \\ & + \dots + \frac{1}{2^{n-1}} = 1 + \frac{1 - \frac{1}{2^n}}{1 - \frac{1}{2}} < 1 + \frac{1}{1 - \frac{1}{2}} = 3. \end{aligned}$$

Итак, последовательность $\left\{\left(1 + \frac{1}{n}\right)^n\right\}$ — монотонно возрастающая и ограниченная сверху, следовательно, она имеет предел. Этот предел принято обозначать буквой e .

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Из неравенства $\left(1 + \frac{1}{n}\right)^n < 3$ следует, что $e \leq 3$. Отбрасывая в равенстве (1) все члены, начиная с четвертого, имеем:

$$\left(1 + \frac{1}{n}\right)^n > 2 + \frac{1}{2} \left(1 - \frac{1}{n}\right)$$

и, переходя к пределу в этом неравенстве, получим

$$e \geq 2 + \frac{1}{2} = 2,5.$$

Таким образом, число e заключено между 2,5 и 3. О более точном вычислении числа e будет сказано дальше, при рассмотрении формулы Тейлора. Можно доказать, что число e — иррациональное и значение $e = 2,71828 \dots$ Заметим еще, что из теоремы о пределе монотонной последовательности следует так же, что неравенство $\left(1 + \frac{1}{n}\right)^n \leq e$ выполняется при любом n . Логарифмы чисел по основанию e называют *натуральными логарифмами* и обозначают $\ln x$. В теоретических исследованиях выбор числа e в качестве основания логарифмов удобен. В практических вычислениях иногда удобнее десятичные логарифмы. Установим связь между логарифмами с различными основаниями.

Возьмем за основания логарифмов положительные числа a и b , $a \neq 1$, $b \neq 1$. Пусть $y = \log_a x$, тогда $a^y = x$.

Логарифмируя это равенство по основанию b , получим:

$$y \log_b a = \log_b x$$

или

$$\log_b x = \log_b a \cdot \log_a x.$$

В частности, при $x = b$

$$\log_b a \cdot \log_a b = 1.$$

Обозначая через $M = \log_b a$ модуль перехода от одного основания к другому, получим:

$$\log_b x = M \log_a x.$$

Полагая $a = e$, $b = 10$, заключаем, что связь между десятичными и натуральными логарифмами дается равенством

$$\lg x = M \ln x,$$

где $M = \lg e \approx 0,43429$.

Раздел 3

ПРЕДЕЛ ФУНКЦИИ

§ 12. Определение предела функции

Пусть функция $f(x)$ определена в некоторой окрестности точки a , т. е. в некотором интервале, содержащем точку a , кроме, быть может, самой точки a . В точке a функция может быть и не определена.

Число A называется *пределом функции* $f(x)$ при x , стремящемся к a (или в точке a), если для любого положительного числа ε найдется такое

положительное число δ , что для всех x , удовлетворяющих условию

$$0 < |x - a| < \delta,$$

выполняется неравенство

$$|f(x) - A| < \varepsilon,$$

или, что то же: если для любого $\varepsilon > 0$ найдется $\delta > 0$, такое, что для всех x , удовлетворяющих условию

$$a - \delta < x < a + \delta, \quad x \neq a,$$

выполняется условие

$$A - \varepsilon < f(x) < A + \varepsilon.$$

Это можно следующим образом иллюстрировать геометрически: возьмем ε -окрестность точки A — интервал $(A - \varepsilon, A + \varepsilon)$ оси OY (рис. 151) и будем рассматривать все значения x из окрестности точки a , для которых соответствующие значения функции не выходят из ε -окрестности точки A .

Из определения следует, что для любого $\varepsilon > 0$ (подчеркнем, что положительное число ε может быть сколь угодно малым) существует $\delta > 0$, такое, что для всех точек симметричного относительно a интервала $(a - \delta, a + \delta)$, $x \neq a$, значения функции не выходят из ε -окрестности точки A . Ясно, что δ зависит от ε , $\delta = \delta(\varepsilon)$, чем больше ε , тем δ , вообще говоря, больше.

Если рассмотреть график функции, изображенный на рис. 152, то станет ясно, что эта функция не имеет предела при $x \rightarrow a$.

Тот факт, что предел функции при $x \rightarrow a$ равен A , записывают так: $\lim_{x \rightarrow a} f(x) = A$ или $f(x) \rightarrow A$ при $x \rightarrow a$.

Только что рассмотренное определение предела функции в точке a будем называть первым определением или $(\varepsilon - \delta)$ — определением.

Пример 1. Докажем, что функция $f(x) = 3x + 2$ при $x \rightarrow 1$ имеет предел, равный 5.

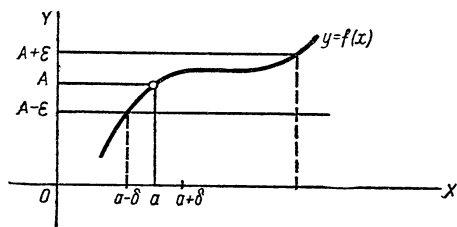


Рис. 151.

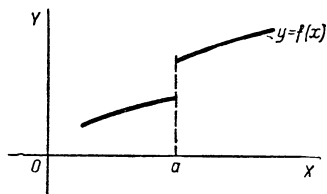


Рис. 152.

Для доказательства установим, что, какое бы положительное число ε мы ни взяли, можно найти такое положительное число δ , что для всех значений x , отличающихся по абсолютной величине от числа 1 меньше чем на δ , будет справедливо неравенство

$$|(3x+2)-5|<\varepsilon.$$

Последнее неравенство можно записать, в виде $3|x-1|<\varepsilon$ или, что то же, $|x-1|<\frac{\varepsilon}{3}$. Поэтому, если взять δ , удовлетворяющее условию $0<\delta\leq\frac{\varepsilon}{3}$, то $3x+2$ будет отличаться от числа 5 по абсолютной величине меньше чем на ε при $|x-1|<\delta$. А это и означает, что $\lim_{x \rightarrow 1} (3x+2) = 5$.

Пример 2. Докажем, что $\lim_{x \rightarrow 0} x^n = 0$ (n — натуральное число). Пусть ε — произвольное положительное число, найдем соответствующее $\delta > 0$, так чтобы для всех x из δ -окрестности точки $x=0$ выполнялось неравенство $|x^n - 0| < \varepsilon$.

Нетрудно видеть, что неравенство $|x^n - 0| = |x|^n < \varepsilon$ выполняется для x , удовлетворяющих условию $|x| < \sqrt[n]{\varepsilon}$. И следовательно, если взять δ , исходя из условия $0 < \delta \leq \sqrt[n]{\varepsilon}$, то получится, что неравенство $|x^n - 0| < \varepsilon$ выполняется при $|x - 0| < \delta$. Таким образом, $\lim_{x \rightarrow 0} x^n = 0$.

Определение предела функции в точке через предел последовательности. Число A называется пределом функции $f(x)$ при $x \rightarrow a$, если для любой последовательности значений аргумента x : $x_1, x_2, \dots, x_n, \dots$, сходящейся к a , $x_n \neq a$, соответствующая последовательность значений функции $f(x_1), f(x_2), \dots, f(x_n), \dots$ сходится к A .

При этом предполагается, что последовательность $\{x_n\}$ принадлежит области определения функции.

Докажем эквивалентность двух определений предела функции в точке.

Доказательство. Пусть A — предел функции $f(x)$ при $x \rightarrow a$ в смысле первого определения. Возьмем произвольное число $\varepsilon > 0$, ему соответствует $\delta > 0$, такое, что для всех x , удовлетворяющих неравенству $0 < |x - a| < \delta$, выполняется неравенство $|f(x) - A| < \varepsilon$.

Рассмотрим любую последовательность $\{x_n\}$, $x_n \neq a$, сходящуюся к a . Тогда для указанного значения δ ($\delta > 0$), зависящего от ε , существует число N , такое, что при $n > N$ выполняется неравенство $|x_n - a| < \delta$. И следовательно, при $n > N$ будет выполняться неравенство $|f(x_n) - A| < \varepsilon$. Но так как $\varepsilon > 0$ выбиралось произвольно, то это и означает, что $f(x_n) \rightarrow A$. Итак, мы доказали, что для любой последовательности $\{x_n\}$, $x_n \neq a$, сходящейся к a , последовательность $\{f(x_n)\}$ сходится к A .

Обратно, пусть A — предел функции $f(x)$ при $x \rightarrow a$ в смысле второго определения. Докажем, что A — предел $f(x)$ при $x \rightarrow a$ в смысле первого определения.

Доказательство. Предположим, что это утверждение неверно, т. е. число A не является пределом функции $f(x)$ при $x \rightarrow a$ в смысле первого определения. Тогда существует такое число $\varepsilon_0 > 0$, что для каждого $\delta > 0$ найдется число x (хотя бы одно), удовлетворяющее условию $0 < |x - a| < \delta$, для которого выполняется неравенство $|f(x) - A| \geq \varepsilon_0$.

Рассмотрим какую-нибудь последовательность положительных чисел $\delta_1, \delta_2, \dots, \delta_n, \dots$, сходящуюся к нулю, например, $\{\delta_n\} = \{\frac{1}{n}\}$. Тогда для каждого δ_n найдется свое x_n , $x_n \neq a$, такое, что $|x_n - a| < \delta_n$, в то время как $|f(x_n) - A| \geq \varepsilon_0$. Так как $\lim_{n \rightarrow \infty} \delta_n = 0$, то, переходя к пределу в неравенстве $|x_n - a| < \delta_n$ получим, что $\lim_{n \rightarrow \infty} x_n - a = 0$, т. е. $x_n \rightarrow a$. Но в то же время последовательность $\{f(x_n)\}$ не сходится к A (так как при

всех n имеет место неравенство $|f(x_n) - A| \geq \varepsilon_0$). Полученное противоречие доказывает теорему.

Заметим, что если число A является пределом функции $f(x)$ при $x \rightarrow a$, то говорят: функция $f(x)$ при $x \rightarrow a$ имеет *конечный предел* (равный A).

Определение предела функции через предел последовательности, как это будет показано дальше, дает возможность из теорем, полученных для сходящихся последовательностей, получить соответствующие теоремы о пределах функций.

Это определение дает также возможность установить для некоторых функций, что они не имеют предела при определенном значении аргумента. Для этого достаточно или показать, что существует последовательность значений x , сходящаяся к рассматриваемому значению аргумента, для которой соответствующая последовательность значений функции не имеет предела, или показать, что существуют две различные последовательности, сходящиеся к рассматриваемому значению аргумента, для которых соответствующие последовательности значений функции имеют различные пределы.

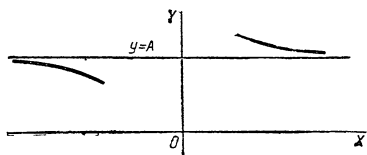


Рис. 153.

Пример. Функция $f(x) = \sin \frac{1}{x}$ не имеет предела при $x \rightarrow 0$.

Действительно, возьмем последовательность $\{x'_n\} = \left\{ \frac{1}{n\pi} \right\}$; $\frac{1}{n\pi} \rightarrow 0$, $f(x'_n) = \sin n\pi = 0 \rightarrow 0$.

Возьмем теперь последовательность $\{x''_n\} = \left\{ \frac{2}{(4n+1)\pi} \right\}$; $\frac{2}{(4n+1)\pi} \rightarrow 0$, при этом $f(x''_n) = \sin(4n+1)\frac{\pi}{2} = \sin \frac{\pi}{2} = 1 \rightarrow 1$.

Функция $\sin \frac{1}{x}$ при $x \rightarrow 0$ не имеет предела, так как по двум различным последовательностям значений x , сходящимся к 0, получаем различные пределы последовательностей значений функции. График функции $f(x) = \sin \frac{1}{x}$ изображен на рис. 168.

Предел функции при стремлении аргумента к бесконечности. Определим предел функции при стремлении аргумента к бесконечности. Число A называется пределом функции $f(x)$ при $x \rightarrow \infty$, если для любого $\varepsilon > 0$ найдется такое число $M > 0$, что для всех значений x , $|x| > M$, выполняется неравенство $|A - f(x)| < \varepsilon$ [при этом предполагается, что функция $f(x)$ определена в окрестности ∞]. Тот факт, что предел функции $f(x)$ при $x \rightarrow \infty$ равен A , записывают таким образом: $\lim_{x \rightarrow \infty} f(x) = A$ или $f(x) \rightarrow A$

при $x \rightarrow \infty$. Приведенное определение предела функции $f(x)$ при $x \rightarrow \infty$ можно сформулировать еще и так: число A называется пределом функции $f(x)$ при $x \rightarrow \infty$, если для любого $\varepsilon > 0$ найдется такая окрестность символа ∞ , что все значения функции $f(x)$ в точках этой окрестности отличаются от числа A по абсолютной величине меньше чем на ε . Если $\lim_{x \rightarrow \infty} f(x) = A$, то график функции $f(x)$ асимптотически приближается к прямой $y = A$ при стремлении x к бесконечности как по положительным, так и по отрицательным значениям (рис. 153).

Аналогично определяются: $\lim_{x \rightarrow +\infty} f(x) = A$ и $\lim_{x \rightarrow -\infty} f(x) = A$. В первом случае для любого числа $\varepsilon > 0$ найдется такое M , что для всех x , $x > M$, выполняется неравенство $|A - f(x)| < \varepsilon$, во втором случае для любого числа $\varepsilon > 0$ найдется такое M , что для всех x , $x < -M$, выполняется неравенство $|A - f(x)| < \varepsilon$.

Графики функций $f(x)$ для каждого из указанных выше случаев изображены на рис. 154, 155.

Определение предела функции $f(x)$ при стремлении аргумента к бесконечности можно также дать следующим образом. Число A называется пределом функции $f(x)$ при $x \rightarrow a$ (a — один из символов ∞ , $+\infty$, $-\infty$), если для любой последовательности $\{x_n\}$, $x_n \rightarrow a$, последовательность $\{f(x_n)\}$ сходится к A .

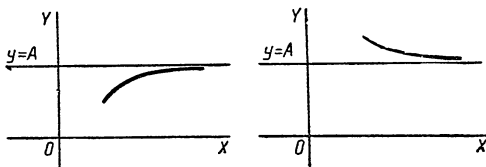


Рис. 154.

Эквивалентность двух рассмотренных определений предела функции при стремлении аргумента к бесконечности доказывается аналогично тому, как это делалось при $x \rightarrow a$, где a — число.

Замечание, приведенное в начале стр. 176, остается в силе и при $x \rightarrow a$, когда a — один из символов ∞ , $+\infty$, $-\infty$.

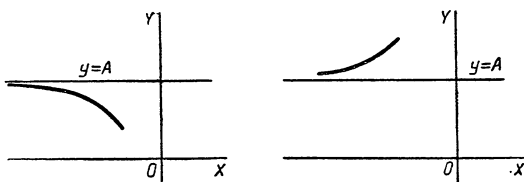


Рис. 155.

§ 13. Основные теоремы о пределах функций

Пусть a — число или один из символов ∞ , $+\infty$, $-\infty$.

Теорема 1. Предел постоянного равен этому постоянному:

$$\lim_{x \rightarrow a} C = C.$$

Действительно, если $f(x)$ в окрестности a сохраняет постоянное значение, равное C , т. е. $f(x) = C$, то для любой последовательности $\{x_n\}$, $x_n \rightarrow a$ ($x_n \neq a$, если a — число) будем иметь $f(x_n) = C \rightarrow C$ и, следовательно, $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} C = C$.

Если функции $f(x)$ и $\varphi(x)$ при $x \rightarrow a$ имеют конечные пределы, то справедливы следующие теоремы:

Теорема 2. $\lim_{x \rightarrow a} [f(x) \pm \varphi(x)] = \lim_{x \rightarrow a} f(x) \pm \lim_{x \rightarrow a} \varphi(x)$.

Теорема 3. $\lim_{x \rightarrow a} [f(x) \varphi(x)] = \lim_{x \rightarrow a} f(x) \lim_{x \rightarrow a} \varphi(x)$.

Теорема 4. $\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \frac{\lim_{x \rightarrow a} f(x)}{\lim_{x \rightarrow a} \varphi(x)}$ при условии, что $\lim_{x \rightarrow a} \varphi(x) \neq 0$.

Заметим, что теоремы 2 и 3 верны также и в случае любого фиксированного числа функций, имеющих при $x \rightarrow a$ конечные пределы.

Эти теоремы доказываются на основании соответствующих теорем о последовательностях.

Докажем, например, теорему 4.

Доказательство. Для любой последовательности $\{x_n\}$, $x_n \neq a$, предел которой равен a , по условию $f(x_n) \rightarrow A$, $\varphi(x_n) \rightarrow B$, $B \neq 0$.

По теореме о частном сходящихся последовательностей имеем, что последовательность $\left\{ \frac{f(x_n)}{\varphi(x_n)} \right\}$ сходится к $\frac{A}{B}$. Но так как последователь-

ность $\{x_n\}$, $x_n \neq a$, любая, удовлетворяющая условию $x_n \rightarrow a$, то, следовательно,

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \frac{A}{B} = \frac{\lim_{x \rightarrow a} f(x)}{\lim_{x \rightarrow a} \varphi(x)}.$$

Как следствие теоремы 3 получаем: $\lim_{x \rightarrow a} C f(x) = C \lim_{x \rightarrow a} f(x)$, т. е. постоянный множитель можно выносить за знак предела.

Пример 1. Пусть $P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$ — многочлен. Покажем, что $\lim_{x \rightarrow a} P(x) = P(a)$, где a — любое число.

Действительно, $\lim_{x \rightarrow a} x = a$ (так как для любой последовательности $\{x_n\}$, $x_n \rightarrow a$, $f(x_n) = x_n \rightarrow a$). Применяя теоремы о пределе произведения и суммы функций, получим:

$$\begin{aligned} \lim_{x \rightarrow a} x^n &= \lim_{x \rightarrow a} x \cdot \lim_{x \rightarrow a} x \cdot \dots \lim_{x \rightarrow a} x = (\lim_{x \rightarrow a} x)^n = a^n, \\ \lim_{x \rightarrow a} P(x) &= \lim_{x \rightarrow a} (a_0 x^n + a_1 x^{n-1} + \dots + a_n) = a_0 (\lim_{x \rightarrow a} x)^n + a_1 (\lim_{x \rightarrow a} x)^{n-1} + \\ &\quad + \dots + a_n = P(\lim_{x \rightarrow a} x) = P(a). \end{aligned}$$

Таким образом, $\lim_{x \rightarrow a} P(x) = P(a)$, в частности $\lim_{x \rightarrow -1} (x^3 + 5x + 1) = (-1)^3 + 5(-1) + 1 = -5$.

Пример 2. Пусть $R(x) = \frac{P(x)}{Q(x)}$ — рациональная функция, a — любое число и $Q(a) \neq 0$. Покажем, что $\lim_{x \rightarrow a} \frac{P(x)}{Q(x)} = \frac{P(a)}{Q(a)}$.

В самом деле, как следует из примера 1, $\lim_{x \rightarrow a} P(x) = P(a)$ и $\lim_{x \rightarrow a} Q(x) = Q(a)$. По теореме 4 имеем $\lim_{x \rightarrow a} \frac{P(x)}{Q(x)} = \frac{\lim_{x \rightarrow a} P(x)}{\lim_{x \rightarrow a} Q(x)} = \frac{P(a)}{Q(a)}$.¹⁾

Условимся говорить, что некоторый факт имеет место *вблизи* a , если этот факт имеет место для всех x некоторой окрестности a , кроме a .

Теорема 5. Если $f(x) > 0$ (≥ 0) вблизи a и $\lim_{x \rightarrow a} f(x) = A$, то $A \geq 0$.

Теорема также верна, если знак $> (\geq)$ заменить знаком $< (\leq)$.

Теорема 6. Если вблизи a выполняются неравенства

$$\varphi(x) \leq f(x) \leq \psi(x)$$

и

$$\lim_{x \rightarrow a} \varphi(x) = \lim_{x \rightarrow a} \psi(x) = A,$$

то и $\lim_{x \rightarrow a} f(x) = A$.

Теоремы 5 и 6 доказываются на основании соответствующих теорем для последовательностей. Читателю предлагается самостоятельно провести соответствующие рассуждения.

Функция $f(x)$ называется *ограниченной вблизи* a , если существует число $M > 0$, такое, что неравенство $|f(x)| < M$ выполняется вблизи a .

Теорема 7. Если функция $f(x)$ при $x \rightarrow a$ имеет конечный предел, то она ограничена вблизи a .

¹⁾ В примерах 1 и 2 доказано следующее: для рациональной функции $f(x)$, целой и дробной, при условии, что ее знаменатель отличен от нуля в точке a , справедливо равенство

$$\lim_{x \rightarrow a} f(x) = f(\lim_{x \rightarrow a} x) = f(a).$$

Это равенство, заключающееся в возможности перестановки знаков предела и функции, выражает свойство непрерывности рассматриваемых функций в точке a (см. § 19).

Пусть $\lim_{x \rightarrow a} f(x) = A$. Возьмем $\epsilon = 1$, тогда вблизи a выполняется неравенство $|f(x) - A| < 1$. Так как $|f(x)| = |[f(x) - A] + A| \leq |f(x) - A| + |A|$, то вблизи a справедливо неравенство

$$|f(x)| < 1 + |A|,$$

или, что то же, неравенство $|f(x)| < M$, где $M = 1 + |A|$. Следовательно, функция $f(x)$ ограничена вблизи a .

§ 14. Бесконечно малые функции и их свойства

Функция $f(x)$ называется *бесконечно малой* при $x \rightarrow a$ (a — число или один из символов ∞ , $+\infty$, $-\infty$), если $\lim_{x \rightarrow a} f(x) = 0$.

Пример 1. $f(x) = x^n$ (n — натуральное число) — бесконечно малая при $x \rightarrow 0$, так как $\lim_{x \rightarrow 0} x^n = 0$ (см. пример 2, § 12)¹⁾.

Пример 2. Функция $f(x) = \sin x$ является бесконечно малой при $x \rightarrow 0$, так как $\lim_{x \rightarrow 0} \sin x = 0$ ²⁾.

Теорема. Для того чтобы функция $f(x)$ при $x \rightarrow a$ имела предел, равный A , необходимо и достаточно, чтобы вблизи a выполнялось равенство $f(x) = A + \alpha(x)$, где $\alpha(x) \rightarrow 0$ при $x \rightarrow a$.

Доказательство. Пусть $f(x) - A = \varphi(x)$, тогда $f(x) = A + \varphi(x)$. Если $\lim_{x \rightarrow a} f(x) = A$, то для любой последовательности $\{x_n\}$, $x_n \neq a$, предел которой равен a , $f(x_n) \rightarrow A$. Но $f(x_n) = A + \varphi(x_n)$, откуда, учитывая аналогичную теорему о последовательностях, получим $\varphi(x_n) \rightarrow 0$. Следовательно, $\varphi(x) \rightarrow 0$ при $x \rightarrow a$, т. е. $\varphi(x)$ — бесконечно малая при $x \rightarrow a$. Обратно, если $f(x) = A + \alpha(x)$, где $\alpha(x) \rightarrow 0$ при $x \rightarrow a$, то для любой последовательности $\{x_n\}$, $x_n \neq a$, предел которой равен a , имеем $f(x_n) = A + \alpha(x_n)$, причем $\alpha(x_n) \rightarrow 0$. Следовательно, $f(x_n) \rightarrow A$ и, значит, $\lim_{x \rightarrow a} f(x) = A$.

Теоремы о бесконечно малых.

Теорема 1. Сумма фиксированного числа бесконечно малых при $x \rightarrow a$ функций есть функция бесконечно малая при $x \rightarrow a$.

Теорема 2. Произведение фиксированного числа бесконечно малых при $x \rightarrow a$ функций есть функция бесконечно малая при $x \rightarrow a$.

Теоремы 1 и 2 получаются из соответствующих теорем о пределах функций.

Теорема 3. Произведение функции бесконечно малой при $x \rightarrow a$ на функцию, ограниченную вблизи a , есть функция бесконечно малая при $x \rightarrow a$.

Эту теорему можно доказывать на основании соответствующей теоремы для последовательностей.

Пример. Функция $\sin \frac{1}{x}$ является ограниченной вблизи $x = 0$, так как $|\sin \frac{1}{x}| \leq 1$ при $x \neq 0$, x — бесконечно малая при $x \rightarrow 0$, следовательно, $x \sin \frac{1}{x}$ — бесконечно малая при $x \rightarrow 0$.

¹⁾ Заметим, что x^n (n — натуральное число) при $x \rightarrow 1$ не является бесконечно малой, так как $\lim_{x \rightarrow 1} x^n = 1$. Поэтому нельзя говорить, что $f(x)$ — бесконечно малая, не указав, к какому значению стремится аргумент x , естественно, кроме случая, когда функция $f(x)$ тождественно равна нулю.

²⁾ Равенство $\lim_{x \rightarrow 0} \sin x = 0$, означающее, как будет указано ниже, непрерывность функции $\sin x$ в точке $x = 0$, можно доказать следующим образом. Для всех x из интервала $(-\frac{\pi}{2}, \frac{\pi}{2})$ $\sin x$ удовлетворяет неравенству $|\sin x| \leq |x|$ (см. рис. 159). Переходя в этом неравенстве к пределу при $x \rightarrow 0$, получим $\sin x \rightarrow 0$.

Как следствие теоремы 3 получаем: произведение функции, имеющей конечный предел при $x \rightarrow a$, на функцию бесконечно малую при $x \rightarrow a$ является функцией бесконечно малой при $x \rightarrow a$.

Например: $P(x) \sin x$ ($P(x)$ — многочлен) — бесконечно малая при $x \rightarrow 0$.

§ 15. Бесконечно большие функции и их связь с бесконечно малыми

Нам уже известно определение предела (конечного предела) функции $f(x)$ при $x \rightarrow a$ (a — число или один из символов ∞ , $+\infty$, $-\infty$), а также известны основные теоремы о пределах. Определим теперь *бесконечный предел функции* $f(x)$ при $x \rightarrow a$ (a — число).

Предел функции $f(x)$ при $x \rightarrow a$ (a — число) равен бесконечности, если для любого числа $M > 0$ существует число $\delta > 0$, такое, что неравенство $|f(x)| > M$ выполняется для всех x , удовлетворяющих условию

$$0 < |x - a| < \delta.$$

Тот факт, что предел функции $f(x)$ при $x \rightarrow a$ равен бесконечности, обозначают так: $\lim_{x \rightarrow a} f(x) = \infty$ или $f(x) \rightarrow \infty$ при $x \rightarrow a$.

Предел функции $f(x)$ при $x \rightarrow a$ равен $+\infty$ ($-\infty$), если для любого числа M существует число $\delta > 0$, такое, что $f(x) > M$ ($f(x) < M$) для всех x , удовлетворяющих условию $0 < |x - a| < \delta$.

Это соответственно, обозначают следующим образом:

$$\begin{aligned} \lim_{x \rightarrow a} f(x) &= +\infty \text{ или } f(x) \rightarrow +\infty \text{ при } x \rightarrow a, \\ \lim_{x \rightarrow a} f(x) &= -\infty \text{ или } f(x) \rightarrow -\infty \text{ при } x \rightarrow a. \end{aligned}$$

Графики таких функций изображены на рис. 156.

Определение бесконечного предела функции при стремлении аргумента к бесконечности. Предел функции $f(x)$ при $x \rightarrow \infty$ равен бесконечности (обозначают $\lim_{x \rightarrow \infty} f(x) = \infty$), если для любого числа $M > 0$ найдется число $K > 0$, такое, что для всех x , удовлетворяющих неравенству $|x| > K$, выполняется неравенство $|f(x)| > M$.

Предлагается самостоятельно сформулировать определение предела функции $f(x)$, равного $+\infty$ ($-\infty$) при $x \rightarrow a$, если a — один из символов ∞ , $+\infty$, $-\infty$.

Определение бесконечного предела функции через последовательность дается так же, как и в случае конечного предела: $\lim_{x \rightarrow a} f(x) = A$ (a — число или один из символов ∞ , $+\infty$, $-\infty$ и A — один из символов ∞ , $+\infty$, $-\infty$), если для любой последовательности $\{x_n\}$ ($x_n \neq a$, если a — число) предел которой равен a , последовательность $\{f(x_n)\}$ удовлетворяет условию $f(x_n) \rightarrow A$.

Доказательства эквивалентности двух определений мы не приводим. Функция $f(x)$ называется *бесконечно большой* при $x \rightarrow a$ (a — число или один из символов ∞ , $+\infty$, $-\infty$), если

$$\lim_{x \rightarrow a} f(x) = A,$$

где A — один из символов ∞ , $+\infty$, $-\infty$.

Связь между функциями бесконечно малыми и бесконечно большими при $x \rightarrow a$ устанавливается **теоремой**: если функция $f(x)$ — бесконечно

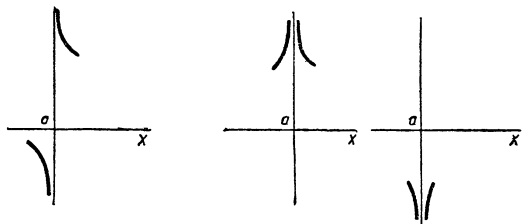


Рис. 156.

малая при $x \rightarrow a$ и $f(x) \neq 0$ вблизи a , то функция $\frac{1}{f(x)}$ — бесконечно большая при $x \rightarrow a$ и, обратно, если функция $f(x)$ — бесконечно большая при $x \rightarrow a$, то функция $\frac{1}{f(x)}$ — бесконечно малая при $x \rightarrow a$ (доказательство следует из аналогичной теоремы для последовательностей).

Пример 1. $f(x) = x$ — бесконечно малая функция при $x \rightarrow 0$;

$$f(x) = \frac{1}{x} \text{ — бесконечно большая функция при } x \rightarrow 0;$$

$$f(x) = x \text{ — бесконечно большая функция при } x \rightarrow \infty;$$

$$f(x) = \frac{1}{x} \text{ — бесконечно малая функция при } x \rightarrow \infty;$$

(См. график функции $f(x) = \frac{1}{x}$ на рис. 157).

Пример 2. $\lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$ и $\lim_{x \rightarrow \infty} \frac{1}{x} \sin \frac{1}{x} = 0$ (как произведение ограниченной функции на бесконечно малую).

Теорема о пределе сложной функции. Пусть $u = \varphi(x)$, $y = f(u)$, причем область изменения первой функции входит в область определения второй. Тогда можно рассматривать сложную функцию $y = f[\varphi(x)]$. Область определения этой сложной функции совпадает с областью определения функции $u = \varphi(x)$.

Теорема. Если $\lim_{x \rightarrow a} \varphi(x) = b$, $\lim_{u \rightarrow b} f(u) = A$ и, кроме того, $A = f(b)$, то сложная функция $y = f[\varphi(x)]$ имеет предел при $x \rightarrow a$, равный A . (Здесь a — число или один из символов ∞ , $+\infty$, $-\infty$, A и b — числа.)

Доказательство. По условию $\varphi(x) \rightarrow b$ при $x \rightarrow a$, следовательно, для любой последовательности $\{x_n\}$, $x_n \neq a$, предел которой равен a , последовательность $\{\varphi(x_n)\}$ имеет предел, равный b (при этом предполагается, что последовательность $\{x_n\}$ взята из области определения функции $u = \varphi(x)$).

Так как по условию еще $f(u) \rightarrow f(b)$ при $u \rightarrow b$, то для любой последовательности $\{u_n\}$, $u_n \neq b$, из области определения функции $y = f(u)$, предел которой равен b , выполняется условие $f(u_n) \rightarrow A = f(b)$.

Это условие выполняется и для последовательности $\{u_n\}$, где $u_n = \varphi(x_n)$. Надобность в требовании $\varphi(x_n) \neq b$ отпадает¹⁾, так как по условию $A = f(b)$.

Таким образом, для любой последовательности $\{x_n\}$, предел которой равен a , последовательность $\{f[\varphi(x_n)]\}$ имеет пределом $A = f(b)$, а это означает, что

$$\lim_{x \rightarrow a} f[\varphi(x)] = A.$$

Теорема о пределе сложной функции дает возможность производить замену переменных в операции перехода к пределу. Пусть требуется найти $\lim_{x \rightarrow a} f(x)$. Предположим, что $f(x)$ можно записать в виде $f(x) = F[\varphi(x)]$, причем $u = \varphi(x)$ при $x \rightarrow a$ имеет предел, который обозначим через b , и $F(u)$ имеет при $u \rightarrow b$ предел, равный $F(b)$, тогда

$$\lim_{x \rightarrow a} f(x) = \lim_{u \rightarrow b} F(u) = F(b).$$

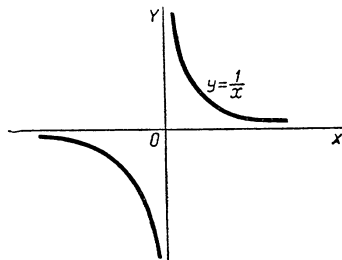


Рис. 157.

¹⁾ Учитывается, что если в какую-нибудь сходящуюся последовательность любым образом вставить любое количество новых членов, равных ее пределу, то новая последовательность будет сходиться к тому же пределу.

Пример 1. Найти $\lim_{x \rightarrow 0} \sin x^2$.

Решение. Положим $u = x^2$. Тогда $u \rightarrow 0$ при $x \rightarrow 0$ и $\sin u \rightarrow \sin 0 = 0$ при $u \rightarrow 0$.

Следовательно, $\lim_{x \rightarrow 0} \sin x^2 = \lim_{u \rightarrow 0} \sin u = 0$.

Пример 2. Найти $\lim_{x \rightarrow 0} \ln \frac{\sin x}{x}$.

Решение. Положим $u = \frac{\sin x}{x}$, тогда $u \rightarrow 1$ при $x \rightarrow 0$ (см. § 17). При $u \rightarrow 1$ функция $\ln u \rightarrow \ln 1 = 0$ ¹⁾.

Следовательно, сложная функция $\ln \frac{\sin x}{x}$ при $x \rightarrow 0$ имеет предел, равный нулю:

$$\lim_{x \rightarrow 0} \ln \frac{\sin x}{x} = 0.$$

§ 16. Односторонние пределы функции

Число A называется *пределом справа функции $f(x)$ при $x \rightarrow a$* (в точке a), если для любого числа $\varepsilon > 0$ найдется число $\delta > 0$, такое, что неравенство $|A - f(x)| < \varepsilon$ выполняется для всех x , удовлетворяющих условию $0 < x - a < \delta$.

Аналогично определяется *предел слева функции $f(x)$ при $x \rightarrow a$* , если неравенство $|f(x) - A| < \varepsilon$ выполняется для всех x , для которых $0 < a - x < \delta$ (см. рис. 158).

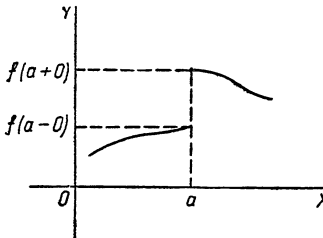


Рис. 158.

Предел справа в точке a обозначается через $f(a+0)$, $\lim_{x \rightarrow a} f(x)$ или $\lim_{x \rightarrow a+0} f(x)$.

Предел слева в точке a обозначается $f(a-0)$, $\lim_{x \rightarrow a} f(x)$ или $\lim_{x \rightarrow a-0} f(x)$.

Если $a=0$, то обозначаются пределы справа и слева соответственно через $f(+0)$, $f(-0)$.

Если функция $f(x)$ имеет предел при $x \rightarrow a$, то, как следует из определения предела, существуют и односторонние пределы функции $f(x)$ при $x \rightarrow a$ и $f(a+0) = f(a-0) = \lim_{x \rightarrow a} f(x)$.

Обратно, если односторонние пределы функции $f(x)$ в точке a существуют и равны $f(a+0) = f(a-0)$, то $\lim_{x \rightarrow a} f(x)$ существует и $\lim_{x \rightarrow a} f(x) = f(a+0) = f(a-0)$.

Действительно, обозначим $f(a+0) = f(a-0) = A$, тогда для любого числа $\varepsilon > 0$ найдется число $\delta_1 > 0$, такое, что неравенство $|A - f(x)| < \varepsilon$ выполняется для всех x , удовлетворяющих условию $0 < x - a < \delta_1$, и число $\delta_2 > 0$, такое, что неравенство $|A - f(x)| < \varepsilon$ выполняется для всех x , удовлетворяющих условию $0 < a - x < \delta_2$. Следовательно, для всех x , удовлетворяющих условию $0 < |x - a| < \delta$, где δ — наименьшее из чисел δ_1 и δ_2 , имеет место неравенство $|A - f(x)| < \varepsilon$, т. е. $f(x) \rightarrow A$ при $x \rightarrow a$.

¹⁾ Для нахождения предела функции $\ln u$ при $u \rightarrow 1$ мы воспользовались равенством

$$\lim_{u \rightarrow 1} \ln u = \ln 1 = 0.$$

Заметим, что и в дальнейшем, при рассмотрении ряда примеров, иллюстрирующих теорию, мы будем пользоваться равенством $\lim_{x \rightarrow a} f(x) = f(\lim_{x \rightarrow a} x) = f(a)$, которое, чтобы не усложнять изложение, не всегда будем доказывать. Ниже, в разделе IV этой главы, где говорится о непрерывности функции, это равенство будет рассмотрено более подробно. Следует иметь в виду, что это замечание относится только к некоторым примерам и никак не влияет на изложение теории.

Определение бесконечных пределов справа (слева) функции $f(x)$ при $x \rightarrow a$ предлагается сформулировать самостоятельно.

Пример 1. $f(x) = \frac{1}{x}$. В этом случае $f(+0) = +\infty$, $f(-0) = -\infty$.

Пример 2. $f(x) = \ln x$. Здесь $f(+0) = -\infty$, а $f(-0)$ определить нельзя, так как функция $\ln x$ не определена для отрицательных значений x .

Пример 3. $f(x) = e^{\frac{1}{x}}$. В этом случае $f(+0) = +\infty$, $f(-0) = 0$. График функции $f(x) = e^{\frac{1}{x}}$ изображен на рис. 167.

Определение одностороннего предела с помощью предела числовой последовательности дается так же, как и предела при $x \rightarrow a$, с той лишь разницей, что теперь рассматриваются $x_n > a$ при всех n в случае правого предела и $x_n < a$ при всех n в случае левого предела.

Заметим, что все теоремы о пределах функций остаются верными и в случае односторонних пределов.

§ 17. Некоторые замечательные пределы

1. Предел отношения многочленов при стремлении аргумента к бесконечности. Пусть $P(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$, $Q(x) = b_0x^m + b_1x^{m-1} + \dots + b_m$ ($a_0 \neq 0$; $b_0 \neq 0$). Тогда

$$\begin{aligned} \frac{P(x)}{Q(x)} &= \frac{a_0x^n + a_1x^{n-1} + \dots + a_n}{b_0x^m + b_1x^{m-1} + \dots + b_m} = \frac{x^n \left(a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n} \right)}{x^m \left(b_0 + \frac{b_1}{x} + \dots + \frac{b_m}{x^m} \right)} = \\ &= x^{n-m} \frac{a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n}}{b_0 + \frac{b_1}{x} + \dots + \frac{b_m}{x^m}}. \end{aligned}$$

Но

$$\lim_{x \rightarrow \infty} \frac{a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n}}{b_0 + \frac{b_1}{x} + \dots + \frac{b_m}{x^m}} = \frac{a_0}{b_0}.$$

Следовательно,

$$\lim_{x \rightarrow \infty} \frac{P(x)}{Q(x)} = \begin{cases} 0, & \text{если } n < m, \\ \frac{a_0}{b_0}, & \text{если } n = m, \\ \infty, & \text{если } n > m. \end{cases}$$

Примеры: 1) $\lim_{x \rightarrow \infty} \frac{3x^2 + 7x + 5}{4x^2 + 9} = \frac{3}{4}$.
2) $\lim_{x \rightarrow \infty} \frac{x^3 + 3}{3x^4 + 5x - 3} = 0$.

2. Предел отношения $\frac{\sin x}{x}$ при $x \rightarrow 0$. Функция $f(x) = \frac{\sin x}{x}$ — четная, так как

$$f(-x) = \frac{\sin(-x)}{-x} = \frac{\sin x}{x} = f(x).$$

По условию $x \rightarrow 0$, следовательно, достаточно рассмотреть значения x , удовлетворяющие неравенствам $0 < x < \frac{\pi}{2}$.

Возьмем дугу окружности радиуса R и угол, радианная мера которого равна x (рис. 159).

$$OA = R, \quad \sin x = \frac{MK}{R}, \quad \operatorname{tg} x = \frac{AT}{R}.$$

Очевидно, пл. $\triangle OAM < \text{пл. сект. } OAM < \text{пл. } \triangle OAT$ или $\frac{1}{2} OA \cdot MK < \frac{1}{2} OA \cdot \tilde{AM} < \frac{1}{2} OA \cdot AT$ и $\frac{1}{2} R^2 \sin x < \frac{1}{2} R^2 x < \frac{1}{2} R^2 \operatorname{tg} x$, откуда получим $\sin x < x < \operatorname{tg} x$. Разделим эти неравенства на $\sin x$ ($\sin x > 0$). Тогда $1 < \frac{x}{\sin x} < \frac{1}{\cos x}$ и, следовательно, $1 > \frac{\sin x}{x} > \cos x$.

Отношение $\frac{\sin x}{x}$ при $0 < x < \frac{\pi}{2}$ заключено между 1 и $\cos x \rightarrow 1$ при $x \rightarrow 0^+$). По теореме VI § 13 получим:

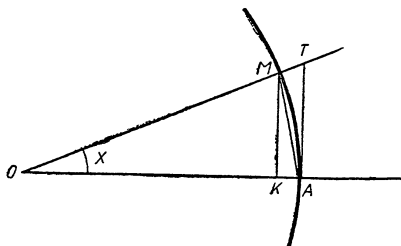


Рис. 159.

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\sin x}{x} = 1.$$

В силу четности функции $\frac{\sin x}{x}$ имеем:

$$\lim_{\substack{x \rightarrow 0 \\ x < 0}} \frac{\sin x}{x} = 1.$$

И окончательно

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

Следует иметь в виду, что предел функции $\frac{\sin x}{x}$ при $x \rightarrow a$, $a \neq 0$, отличен от единицы. Так, например, $\lim_{x \rightarrow \frac{\pi}{2}} \frac{\sin x}{x} = \frac{2}{\pi}$ и $\lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$.

Примеры. 1) $\lim_{x \rightarrow 0} \frac{\operatorname{tg} x}{x} = \lim_{x \rightarrow 0} \frac{\sin x}{x \cos x} = \lim_{x \rightarrow 0} \left(\frac{\sin x}{x} \cdot \frac{1}{\cos x} \right) = 1$.

$$2) \lim_{x \rightarrow 0} \frac{\sin 7x}{\sin 3x} = \lim_{x \rightarrow 0} \frac{7 \frac{\sin 7x}{7x}}{3 \frac{\sin 3x}{3x}} = \frac{7}{3}.$$

3. $\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e$. Как известно, $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$. Докажем, что

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Действительно, пусть $x > 1$. Положим $n = E(x)$, тогда $x = n + \alpha$, где n — натуральное число, а α удовлетворяет условию $0 \leq \alpha < 1$. При $x \rightarrow +\infty$ ($n \rightarrow \infty$) имеют место следующие неравенства:

$$\left(1 + \frac{1}{n+1}\right)^n < \left(1 + \frac{1}{x}\right)^x < \left(1 + \frac{1}{n}\right)^{n+1}.$$

¹⁾ Функция $\cos x$ — четная, поэтому если доказать, что $\lim_{x \rightarrow +0} \cos x = 1$, то тем самым будет доказано, что $\lim_{x \rightarrow 0} \cos x = 1$. Равенство $\lim_{x \rightarrow +0} \cos x = 1$ можно доказать, например, так: для всех x , удовлетворяющих условию $0 < x < \frac{\pi}{2}$, справедливо неравенство $0 < 1 - \cos x = 2 \sin^2 \frac{x}{2} < \frac{x^2}{2}$, откуда, переходя к пределу при $x \rightarrow 0$ ($x > 0$), получим $1 - \cos x \rightarrow 0$, т. е. $\lim_{x \rightarrow +0} \cos x = 1$.

Как нетрудно видеть,

$$\left(1 + \frac{1}{n}\right)^{n+1} = \left(1 + \frac{1}{n}\right)^n \cdot \left(1 + \frac{1}{n}\right) \rightarrow e \cdot 1 = e$$

и

$$\left(1 + \frac{1}{n+1}\right)^n = \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{1 + \frac{1}{n+1}} \rightarrow \frac{e}{1} = e.$$

Откуда по теореме VI § 13 получим:

$$\lim_{x \rightarrow +\infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Пусть теперь $x < -1$, положим $x = -y$. Тогда

$$\begin{aligned} \left(1 + \frac{1}{x}\right)^x &= \left(1 - \frac{1}{y}\right)^{-y} = \left(\frac{y}{y-1}\right)^y = \left(1 + \frac{1}{y-1}\right)^y = \\ &= \left(1 + \frac{1}{y-1}\right)^{y-1} \left(1 + \frac{1}{y-1}\right) \rightarrow e \cdot 1 = e \end{aligned}$$

при $y \rightarrow +\infty$ ($x \rightarrow -\infty$).

Учитывая оба случая, окончательно получим:

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Если в этом равенстве положить $\frac{1}{x} = \alpha$ ($\alpha \rightarrow 0$ при $x \rightarrow \infty$), то оно запишется в виде

$$\lim_{\alpha \rightarrow 0} (1 + \alpha)^{\frac{1}{\alpha}} = e.$$

Пример.

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \lim_{n \rightarrow \infty} \left[\left(1 + \frac{x}{n}\right)^{\frac{n}{x}}\right]^x = e^x.$$

§ 18. Сравнение бесконечно малых

Пусть $\alpha(x)$, $\beta(x)$, $\gamma(x)$ — функции, бесконечно малые при $x \rightarrow a$ (a — число или один из символов ∞ , $+\infty$, $-\infty$). Для краткости условимся, опуская аргумент, обозначать эти функции просто α , β , γ .

Бесконечно малые функции часто сравнивают между собой по быстроте стремления к нулю. Так например, из двух функций $\alpha = x$, $\beta = x^{10}$, бесконечно малых при $x \rightarrow 0$, x^{10} стремится к нулю «быстрее», чем x .

Уточним, какой смысл вкладывается в слово «быстрее».

Если предел отношения $\frac{\alpha}{\beta}$ при $x \rightarrow a$ равен нулю, то говорят, что α — *бесконечно малая более высокого порядка*, чем β (или β — бесконечно малая более низкого порядка, чем α)¹⁾.

Если предел отношения $\frac{\alpha}{\beta}$ при $x \rightarrow a$ конечен и отличен от нуля, то α и β называются *бесконечно малыми одного порядка*. В частном случае, если этот предел равен единице, то бесконечно малые α и β называются *эквивалентными бесконечно малыми* при $x \rightarrow a$. Тот факт, что α и β — эквивалентные бесконечно малые, записывают так: $\alpha \sim \beta$.

¹⁾ При этом предполагается, что $\beta(x)$ в некоторой окрестности a отлична от нуля.

Пример 1. $\alpha = \sin x$, $\beta = 2x$ — функции, бесконечно малые при $x \rightarrow 0$.
Так как

$$\lim_{x \rightarrow 0} \frac{\alpha}{\beta} = \lim_{x \rightarrow 0} \frac{\sin x}{2x} = \frac{1}{2},$$

то α и β — бесконечно малые одного порядка.

Пример 2. $\alpha = \sin x$, $\beta = x$ — бесконечно малые при $x \rightarrow 0$.

$$\lim_{x \rightarrow 0} \frac{\alpha}{\beta} = \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1,$$

т. е. $\sin x \sim x$ при $x \rightarrow 0$.

Аналогично показывается, что $\operatorname{tg} x \sim x$ при $x \rightarrow 0$.

Пример 3. $\alpha = x \sin x$, $\beta = x$.

Так как

$$\lim_{x \rightarrow 0} \frac{\alpha}{\beta} = \lim_{x \rightarrow 0} \sin x = 0,$$

то α — бесконечно малая более высокого порядка, чем β , при $x \rightarrow 0$.

Бесконечно малая α называется бесконечно малой порядка k относительно бесконечно малой β , если предел отношения $\frac{\alpha}{\beta^k}$ при $x \rightarrow a$ конечен и отличен от нуля.

Например, $\alpha = x \sin x$, $\beta = x$ — функции, бесконечно малые при $x \rightarrow 0$. При решении примера 3 установлено, что α — бесконечно малая более высокого порядка, чем β . Теперь мы можем сказать, что α — бесконечно малая второго порядка относительно β , так как

$$\lim_{x \rightarrow 0} \frac{\alpha}{\beta^2} = \lim_{x \rightarrow 0} \frac{x \sin x}{x^2} = 1.$$

Заметим, что не всякие бесконечно малые можно сравнивать между собой. Например, функции $\alpha = x \sin \frac{1}{x}$ и $\beta = x$ — бесконечно малые при $x \rightarrow 0$, а отношение $\frac{\alpha}{\beta} = \sin \frac{1}{x}$ не имеет предела при $x \rightarrow 0$. В этом случае говорят, что бесконечно малые α и β *несравнимы*.

Отметим следующие свойства эквивалентных бесконечно малых:

1) $\alpha \sim \alpha \quad \left(\lim_{x \rightarrow a} \frac{\alpha}{\alpha} = 1 \right).$

2) Если $\alpha \sim \beta$, то $\beta \sim \alpha$.

$$\left(\lim_{x \rightarrow a} \frac{\beta}{\alpha} = \lim_{x \rightarrow a} \frac{1}{\frac{\alpha}{\beta}} = 1 \right).$$

3) Если $\alpha \sim \beta$ и $\beta \sim \gamma$, то $\alpha \sim \gamma$.

$$\left(\lim_{x \rightarrow a} \frac{\alpha}{\gamma} = \lim_{x \rightarrow a} \left(\frac{\alpha}{\beta} \cdot \frac{\beta}{\gamma} \right) = 1 \cdot 1 = 1 \right).$$

Для приложений особенно важно следующее свойство эквивалентных бесконечно малых.

4) Если $\alpha \sim \alpha_1$, $\beta \sim \beta_1$ и $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = K$, то и $\lim_{x \rightarrow a} \frac{\alpha_1}{\beta_1} = K$, т. е. предел отношения бесконечно малых не меняется при замене их эквивалентными бесконечно малыми:

$$\lim_{x \rightarrow a} \frac{\alpha}{\beta} = \lim_{x \rightarrow a} \frac{\alpha_1}{\beta_1}.$$

Действительно,

$$\lim_{x \rightarrow a} \frac{\alpha_1}{\beta_1} = \lim_{x \rightarrow a} \left(\frac{\alpha}{\beta} \cdot \frac{\frac{\alpha_1}{\alpha}}{\frac{\beta_1}{\beta}} \right) = \lim_{x \rightarrow a} \frac{\alpha}{\beta}.$$

Из свойства 4, в частности, следует:

а) если $\alpha \sim \alpha_1$ и $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = K$, то $\lim_{x \rightarrow a} \frac{\alpha_1}{\beta} = \lim_{x \rightarrow a} \frac{\alpha}{\beta}$ (так как $\beta \sim \beta$);

б) если $\beta \sim \beta_1$ и $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = K$, то $\lim_{x \rightarrow a} \frac{\alpha}{\beta_1} = \lim_{x \rightarrow a} \frac{\alpha}{\beta}$ (так как $\alpha \sim \alpha$).

Таким образом, при нахождении предела произведения (частного) бесконечно малый сомножитель числителя (знаменателя) можно заменить эквивалентной этому сомножителю бесконечно малой.

Пример 1. Найти $\lim_{x \rightarrow 0} \frac{\operatorname{tg} 5x}{\sin 7x}$.

Так как $\operatorname{tg} 5x \sim 5x$, $\sin 7x \sim 7x$ при $x \rightarrow 0$, то, заменяя сомножители эквивалентными бесконечно малыми, получим:

$$\lim_{x \rightarrow 0} \frac{\operatorname{tg} 5x}{\sin 7x} = \lim_{x \rightarrow 0} \frac{5x}{7x} = \frac{5}{7}.$$

Пример 2. Найти $\lim_{x \rightarrow 0} \frac{x^3}{1 - \cos x}$.

Так как $1 - \cos x = 2 \sin^2 \frac{x}{2} \sim 2 \left(\frac{x}{2}\right)^2$ при $x \rightarrow 0$, то

$$\lim_{x \rightarrow 0} \frac{x^3}{1 - \cos x} = \lim_{x \rightarrow 0} \frac{x^3}{\frac{x^2}{2}} = \lim_{x \rightarrow 0} 2x = 0.$$

Пример 3. $\lim_{x \rightarrow 0} \frac{\operatorname{tg} x}{\sin x^2} = \lim_{x \rightarrow 0} \frac{x}{x^2} = \infty$.

Если α и β — функции, бесконечно малые при $x \rightarrow a$, причем β — бесконечно малая более высокого порядка, чем α , то $\gamma = \alpha + \beta$ — бесконечно малая, эквивалентная α , так как

$$\lim_{x \rightarrow a} \frac{\gamma}{\alpha} = \lim_{x \rightarrow a} \left(1 + \frac{\beta}{\alpha}\right) = 1.$$

В этом случае говорят, что α — главная часть бесконечно малой γ .

Например, $x^2 + x$ — функция, бесконечно малая при $x \rightarrow 0$, x — главная часть этой бесконечно малой.

Заметим, что функции $f(x)$ и $\varphi(x)$, бесконечно большие при $x \rightarrow a$ (a — число или один из символов ∞ , $+\infty$, $-\infty$), можно сравнивать, аналогичным образом, рассматривая предел отношения $\frac{f(x)}{\varphi(x)}$ при $x \rightarrow a$, но на этом мы останавливаться не будем.

Раздел 4

НЕПРЕРЫВНОСТЬ ФУНКЦИИ

§ 19. Основные определения

Функция $f(x)$, определенная в точке x_0 и некоторой ее окрестности, называется *непрерывной в точке x_0* , если предел функции при $x \rightarrow x_0$ равен значению функции в точке x_0 :

$$\lim_{x \rightarrow x_0} f(x) = f(x_0),$$

или, что то же,

$$\lim_{x \rightarrow x_0} f(x) = f(\lim_{x \rightarrow x_0} x).$$

Таким образом, для функции, непрерывной в точке, знаки функции и предела можно переставлять.

Если функция $f(x)$ определена в окрестности точки x_0 (кроме, быть может, самой точки x_0) и не является непрерывной в этой точке, то ее

называют *разрывной в точке* x_0 . В этом случае точку x_0 называют *точкой разрыва функции* $f(x)$.

Учитывая данные выше определения предела функции в точке, получаем соответственно следующие определения.

1. Функция $f(x)$ называется непрерывной в точке x_0 , если для любого числа $\varepsilon > 0$ найдется такое число $\delta > 0$, что неравенство $|f(x) - f(x_0)| < \varepsilon$ выполняется для всех x , удовлетворяющих неравенству $|x - x_0| < \delta$.

Это так называемое $(\varepsilon - \delta)$ -определение непрерывности функции.

2. Функция $f(x)$ называется непрерывной в точке x_0 , если для любой последовательности $\{x_n\}$ значений аргумента, сходящейся к x_0 , соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к $f(x_0)$.

Из определения непрерывности функции $f(x)$ в точке x_0 и теоремы § 14 получаем: для того чтобы функция $f(x)$ была непрерывной в точке x_0 , необходимо и достаточно, чтобы в некоторой окрестности этой точки имело место равенство

$$f(x) = f(x_0) + \alpha(x),$$

где $\alpha(x)$ — бесконечно малая при $x \rightarrow x_0$. Или, что то же, необходимо и достаточно, чтобы $f(x) - f(x_0) = \alpha(x) \rightarrow 0$ при $x \rightarrow x_0$. Это означает, что

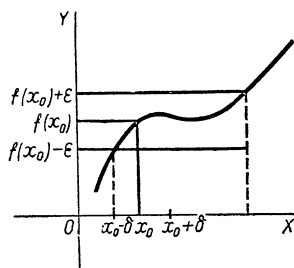


Рис. 160.

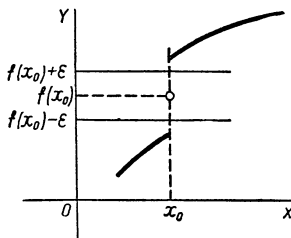


Рис. 161.

для непрерывности функции в точке x_0 необходимо и достаточно, чтобы бесконечно малому изменению аргумента соответствовало бесконечно малое изменение функции.

На рис. 160 изображена функция, непрерывная в точке x_0 .

На рис. 161 изображена функция, имеющая разрыв в точке x_0 . Как видно, здесь не для всякого $\varepsilon > 0$ можно найти соответствующее $\delta > 0$.

Функция $f(x)$ называется *непрерывной справа (слева)* в точке x_0 , если предел функции справа (слева) при $x \rightarrow x_0$ равен значению функции в этой точке. Таким образом, если $f(x_0 + 0) = f(x_0)$, то функция $f(x)$ непрерывна справа в точке x_0 , если же $f(x_0 - 0) = f(x_0)$, то функция $f(x)$ непрерывна слева в точке x_0 .

Заметим, что для односторонней непрерывности функции в точке x_0 вовсе не требуется, чтобы функция была определена в окрестности точки x_0 . Достаточно потребовать, чтобы функция была определена в точке x_0 и справа или слева от нее для значений аргумента, достаточно близких к x_0 (в случае непрерывности справа — в правой полуокрестности точки x_0 $[x_0, x_0 + \delta)$, а в случае непрерывности слева — в левой полуокрестности точки x_0 $(x_0 - \delta, x_0]$). На рис. 162 изображены соответственно непрерывные функции справа и слева в точке x_0 .

Нетрудно видеть, что если функция $f(x)$ непрерывна в точке x_0 , то она непрерывна слева и справа в этой точке. Заметим также, что определение непрерывности функции $f(x)$ в точке x_0 можно записать и так:

$$f(x_0 - 0) = f(x_0) = f(x_0 + 0).$$

Свойства функций, непрерывных в точке. Рассмотрим следующие теоремы.

Теорема 1. Сумма, разность, произведение фиксированного числа функций, непрерывных в некоторой точке, непрерывны в этой точке.

Теорема 2. Частное от деления двух функций, непрерывных в некоторой точке, непрерывно в этой точке при условии, что знаменатель не равен нулю в этой точке.

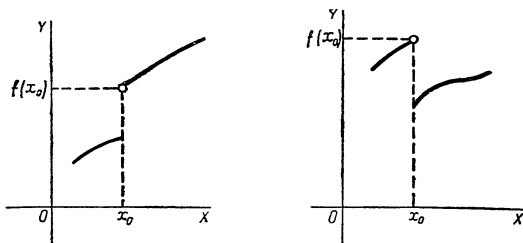


Рис. 162.

Эти теоремы доказываются на основании соответствующих теорем о пределах.

Докажем, например, теорему 2.

Пусть $f(x)$ и $\varphi(x)$ — непрерывные функции в точке x_0 и $\varphi(x_0) \neq 0$. Тогда

$$\lim_{x \rightarrow x_0} \frac{f(x)}{\varphi(x)} = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} \varphi(x)} = \frac{f(x_0)}{\varphi(x_0)},$$

что и требовалось доказать.

Пусть теперь $u = \varphi(x)$, $y = f(u)$, причем область изменения первой функции входит в область определения второй. Как известно, сложная функция $y = f[\varphi(x)]$ будет определена в области определения функции $u = \varphi(x)$.

Теорема 3. Если $u = \varphi(x)$ — непрерывная функция в точке x_0 , $\varphi(x_0) = u_0$, а функция $y = f(u)$ непрерывна в точке u_0 , то функция $y = f[\varphi(x)]$ непрерывна в точке x_0 .

Короче, суперпозиция непрерывных функций есть непрерывная функция.

Доказательство. По условию $\varphi(x)$ — непрерывная функция в точке x_0 , следовательно, $\varphi(x) \rightarrow \varphi(x_0) = u_0$ при $x \rightarrow x_0$.

Функция $f(u)$ непрерывна в точке u_0 , следовательно, $f(u) \rightarrow f(u_0)$ при $u \rightarrow u_0$.

Применяя теорему о пределе сложной функции, получим:

$$\lim_{x \rightarrow x_0} f[\varphi(x)] = \lim_{u \rightarrow u_0} f(u) = f(u_0),$$

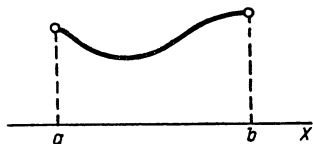


Рис. 163.

или окончательно $\lim_{x \rightarrow x_0} f[\varphi(x)] = f[\varphi(x_0)]$. А это и есть определение непрерывности в точке x_0 .

Определение непрерывности функции на интервале и отрезке. Функция $f(x)$ называется *непрерывной на интервале* (a, b) , если она непрерывна в каждой точке интервала. Функция $f(x)$ называется *непрерывной на отрезке* $[a, b]$, если она непрерывна на интервале (a, b) и, кроме того, непрерывна справа в точке a и непрерывна слева в точке b .

Заметим, что для непрерывности функции на отрезке $[a, b]$ не требуется непрерывность этой функции на концах отрезка. В точках a и b требуется только односторонняя непрерывность функции: для левого конца непрерывность справа, для правого непрерывность слева (рис. 163).

§ 20. Непрерывность некоторых элементарных функций

Пусть $f(x)$ в интервале (a, b) сохраняет постоянное значение $f(x) = C$. Докажем, что $f(x)$ — непрерывная функция на этом интервале.

Действительно, для любого x_0 из интервала (a, b)

$$\lim_{x \rightarrow x_0} f(x) = C = f(x_0).$$

Функция $f(x) = x$ непрерывна для всех x , т. е. на всей числовой оси. В самом деле, для любой точки x_0

$$\lim_{x \rightarrow x_0} f(x) = x_0 = f(x_0).$$

Используя теоремы о непрерывных функциях, получаем непрерывность на всей числовой оси функции $f(x) = x^n$ (n — натуральное число) и многочлена $P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$.

Рациональная функция

$$R(x) = \frac{P(x)}{Q(x)} = \frac{a_0 x^n + a_1 x^{n-1} + \dots + a_n}{b_0 x^m + b_1 x^{m-1} + \dots + b_m}$$

непрерывна для всех значений x , кроме тех, при которых знаменатель равен нулю. Таким образом, функция $R(x)$ непрерывна в ее области определения.

Докажем, что $\sin x$ — функция, непрерывная на всей числовой оси.

Действительно, $|\sin x| \leq |x|$ для $|x| \leq \frac{\pi}{2}$ (см. рис. 159). При $|x| > \frac{\pi}{2}$

также имеем $|\sin x| \leq |x|$, так как $|\sin x| \leq 1$, а $\frac{\pi}{2} > 1$.

Следовательно, для всех значений x справедливо неравенство $|\sin x| \leq |x|$. Так как для произвольной точки x_0

$$\sin x - \sin x_0 = 2 \sin \frac{x - x_0}{2} \cos \frac{x + x_0}{2},$$

то

$$|\sin x - \sin x_0| = 2 \left| \sin \frac{x - x_0}{2} \right| \left| \cos \frac{x + x_0}{2} \right| \leq 2 \left| \frac{x - x_0}{2} \right| \\ |\sin x - \sin x_0| \leq |x - x_0|.$$

Взяв любое $\varepsilon > 0$ и положив $\delta = \varepsilon$, получим, что $|\sin x - \sin x_0| < \varepsilon$ при $|x - x_0| < \delta$. Это и означает непрерывность функции $\sin x$ в точке x_0 .

Аналогично доказывается непрерывность функции $f(x) = \cos x$ в любой точке x_0 . Действительно,

$$\cos x - \cos x_0 = -2 \sin \frac{x + x_0}{2} \sin \frac{x - x_0}{2},$$

$$|\cos x - \cos x_0| \leq 2 \left| \sin \frac{x - x_0}{2} \right| \leq |x - x_0|.$$

И, значит, для любого $\varepsilon > 0$ найдется δ ($\delta = \varepsilon$), такое, что при $|x - x_0| < \delta$ выполняется неравенство $|\cos x - \cos x_0| < \varepsilon$. Таким образом, функция $f(x) = \cos x$ непрерывна в точке x_0 и, следовательно, непрерывна на всей числовой оси.

Из теоремы о непрерывности частного двух непрерывных функций следует, что каждая тригонометрическая функция является непрерывной в ее области определения. Не доказывая, заметим, что любая основная элементарная функция непрерывна в ее области определения.

§ 21. Точки разрыва и их классификация

Пусть функция $f(x)$ определена в некоторой окрестности точки x_0 , кроме, быть может, самой точки x_0 . Из приведенного ранее определения точки разрыва функции получаем, что точка x_0 является точкой разрыва функции $f(x)$, если выполняется одно из условий: или функция $f(x)$ определена в точке x_0 , но не является непрерывной в этой точке, или функция $f(x)$ не определена в точке x_0 .

В первом случае точка x_0 принадлежит области определения функции, во втором случае x_0 не принадлежит области определения функции. Заметим, что для основных элементарных функций возможен только второй случай.

Пусть x_0 — точка разрыва функции $f(x)$. При этом x_0 называется *точкой разрыва первого рода*, если функция $f(x)$ имеет конечные пределы справа и слева в этой точке. Во всех остальных случаях x_0 называется *точкой разрыва второго рода*. Если x_0 — точка разрыва первого рода функции $f(x)$, то график этой функции в точке x_0 может иметь лишь конечный скачок (см. рис. 164). Если же x_0 — точка разрыва второго рода функции $f(x)$, то по крайней мере один из пределов справа или слева в точке x_0 не существует или равен бесконечности.

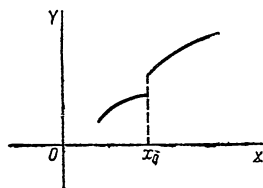


Рис. 164.

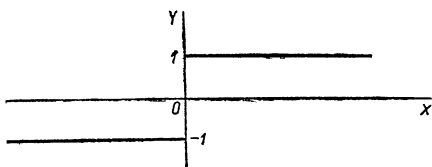


Рис. 165.

Пример 1. $f(x) = \frac{x}{|x|}$
или $f(x) = \begin{cases} 1, & \text{если } x > 0, \\ -1, & \text{если } x < 0 \end{cases}$

[эта функция часто обозначается через $\text{sign } x$ (знак x)]. График функции $f(x)$ изображен на рис. 165. Здесь $x=0$ является точкой разрыва функции $f(x)$, в ней функция неопределена. Так как $f(+0)=1$, $f(-0)=-1$, то, следовательно, при $x=0$ функция имеет разрыв первого рода. Если доопределить $f(x)$ в точке $x=0$, положив $f(0)=1$, то функция $f(x)$ в точке $x=0$ будет также иметь разрыв первого рода, но теперь она уже будет непрерывной справа в этой точке [так как $f(+0)=f(0)$]. Аналогично, если положить $f(0)=-1$, то функция будет непрерывна слева при $x=0$ и, наконец, если положить $f(0)$ равной какому-нибудь числу, отличному от 1 и -1, или $f(x)$ не доопределять в точке $x=0$, то в этой точке функция $f(x)$ не будет непрерывной ни справа, ни слева.

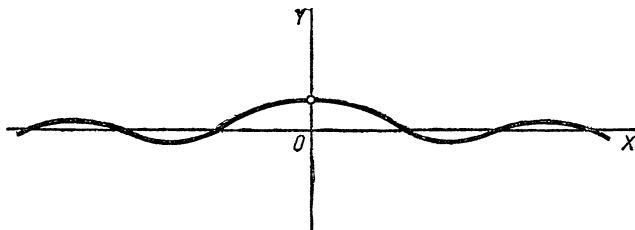


Рис. 166.

Пример 2. $f(x) = E(x)$ (см. ее определение в § 2), график $E(x)$ изображен на рис. 110.

Функция $E(x)$ разрывна в каждой целой точке, т. е. при $x=n$ (n — целое число), так как $f(n+0)=n$, $f(n-0)=n-1$. Все эти точки являются точками разрыва первого рода. В каждой из них функция непрерывна справа.

Пример 3. $f(x) = \frac{\sin x}{x}$ (рис. 166).

В точке $x=0$ функция $\frac{\sin x}{x}$ не определена, но она имеет конечный предел в этой точке: $\lim_{x \rightarrow 0} f(x) = 1$. Следовательно, при $x=0$ функция $\frac{\sin x}{x}$ имеет разрыв

первого рода. Точка $x=0$ в этом случае называется точкой устранимого разрыва. Достаточно доопределить функцию $f(x)$ в точке $x=0$, положив $f(0)=1$, как разрыв устраняется. Итак, функция

$$f(x) = \begin{cases} \frac{\sin x}{x} & \text{при } x \neq 0, \\ 1 & \text{при } x = 0 \end{cases}$$

непрерывна в точке $x=0$. Она непрерывна на всей числовой оси. Заметим, что случай устранимого разрыва есть частный случай разрыва первого рода, когда пределы справа и слева конечны и равны, т. е. существует конечный предел функции в этой точке.

Пример 4. $f(x) = \frac{1}{x}$ (рис. 157) в точке $x=0$ имеет разрыв второго рода. Действительно, здесь $f(+0) = +\infty$, $f(-0) = -\infty$.

Пример 5. Функция $\operatorname{tg} x$ (рис. 130) имеет разрывы второго рода в точках $x = \frac{\pi}{2} + k\pi (k=0, \pm 1, \pm 2, \dots)$.

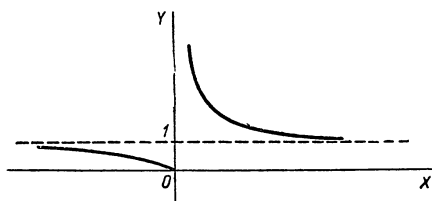


Рис. 167.

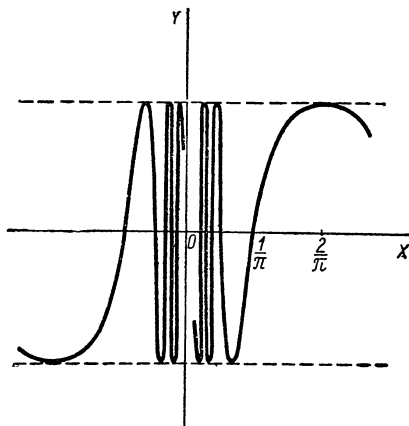


Рис. 168.

Пример 6. $f(x) = e^{\frac{1}{x}}$. Покажем, что $x=0$ — точка разрыва второго рода. Это действительно так, потому что $f(+0) = +\infty$, $f(-0) = 0$. График этой функции изображен на рис. 167.

Пример 7. $f(x) = \sin \frac{1}{x}$. Как было показано, эта функция не имеет предела при $x \rightarrow 0$, следовательно, в точке $x=0$ она имеет разрыв второго рода. Нетрудно видеть, что $\sin \frac{1}{x}$ — нечетная функция; она равна нулю в точках $x = \frac{1}{k\pi} (k = \pm 1, \pm 2, \dots)$, $|\sin \frac{1}{x}| = 1$ при $x = \frac{2}{(2k+1)\pi} (k=0, \pm 1, \pm 2, \dots)$. Кроме того, $\lim_{x \rightarrow \infty} \sin \frac{1}{x} = 0$. График $\sin \frac{1}{x}$ изображен на рис. 168.

Пример 8. Функция Дирихле (см. § 2) разрывна в каждой точке, точки разрыва второго рода.

§ 22. Свойства функций, непрерывных на отрезке

Функция $f(x)$ называется *ограниченной на промежутке*, если существует число $M > 0$, такое, что для всех точек этого промежутка выполняется неравенство

$$|f(x)| < M.$$

Это означает, что график функции на рассматриваемом промежутке не выходит из полосы $-M < y < M$ (рис. 169).

Рассмотрим основные теоремы о свойствах функции, непрерывной на отрезке.

Теорема 1. Функция, непрерывная на отрезке, ограничена на этом отрезке.

При доказательстве теоремы мы воспользуемся известным свойством: функция, непрерывная в точке x_0 , ограничена в некоторой ее окрестности.

Справедливость этого утверждения следует из теоремы 7 § 13.

Доказательство. Допустим, что функция $f(x)$ не ограничена на отрезке $[a, b]$. Разделим отрезок $[a, b]$ пополам, тогда по крайней мере в одном из полученных отрезков эта функция не ограничена (в противном случае она была бы ограничена на отрезке $[a, b]$). Рассмотрим отрезок, на котором функция не ограничена, обозначим его через $[a_1, b_1]$. Далее разделим отрезок $[a_1, b_1]$ пополам и выберем из полученных отрезков тот отрезок (обозначим его $[a_2, b_2]$), на котором функция не ограничена и т. д. Продолжая этот процесс неограниченно, получим последовательность $[a, b], [a_1, b_1], \dots, [a_n, b_n], \dots$ вложенных стягивающихся отрезков, на каждом из которых функция $f(x)$ не ограничена.

По лемме о вложенных отрезках существует точка ξ , принадлежащая всем отрезкам. Функция $f(x)$ непрерывна в точке ξ , следовательно, в некоторой окрестности точки ξ она ограничена¹⁾. В эту окрестность, при достаточно большом n , попадает отрезок $[a_n, b_n]$. Следовательно, на нем функция $f(x)$ будет также ограничена. Последнее противоречит выбору последовательности отрезков. Таким образом, теорема доказана.

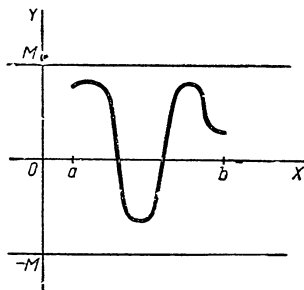


Рис. 169.

Заметим, что теорема перестает быть верной, если функция непрерывна на интервале или полуинтервале. Например, функция $f(x) = \frac{1}{x}$ непрерывна на интервале $(0, 1)$, но не является ограниченной на этом интервале. См. рис. 157.

Теорема 2. Функция $f(x)$, непрерывная на отрезке $[a, b]$, принимает на этом отрезке свое наибольшее и наименьшее значения, т. е. существуют такие точки ξ и η отрезка $[a, b]$, что $f(\xi) = m$, $f(\eta) = M$ и для всех точек этого отрезка выполняются неравенства

$$m \leq f(x) \leq M.$$

Отметим, что свои наибольшее и наименьшее значения функция может принимать и несколько раз. Так, например, $\sin x$ на отрезке $[0, 5\pi]$ несколько раз принимает значение 1 и -1 . Теорема перестает быть верной, если вместо отрезка рассматривать интервал или полуинтервал, например, функция $f(x) = x$ непрерывна на интервале $(0, 1)$, но не принимает на этом интервале ни своего наибольшего, ни наименьшего значения (рис. 170).

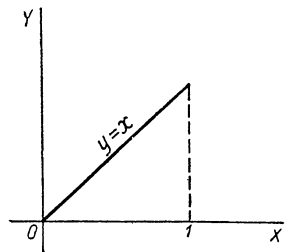


Рис. 170.

Доказательство. Как следует из теоремы 1, функция $f(x)$ ограничена на отрезке $[a, b]$, т. е. существуют числа c и d ($c < d$), такие, что все значения $f(x)$ на отрезке $[a, b]$ лежат внутри отрезка $[c, d]$.

Разделим отрезок $[c, d]$ пополам, тогда имеются две возможности: или все значения функции $f(x)$ принадлежат отрезку $[c, \frac{c+d}{2}]$, или хотя бы одно значение $f(x)$ больше $\frac{c+d}{2}$, т. е. лежит внутри отрезка $[\frac{c+d}{2}, d]$. Если выполняется первое условие, то возьмем отрезок $[c, \frac{c+d}{2}]$, если второе, то возьмем отрезок $[\frac{c+d}{2}, d]$.

¹⁾ Если точка ξ совпадает с одним из концов отрезка $[a, b]$, то рассматривается ее полукрестность.

В любом случае выбранный отрезок обозначим через $[c_1, d_1]$. Аналогично определим отрезок $[c_2, d_2]$ и т. д. Таким образом, получается последовательность вложенных стягивающихся отрезков $[c_1, d_1], [c_2, d_2], \dots, [c_n, d_n], \dots$ обладающих следующим свойством: для любого числа c_n существует хотя бы одно значение функции $f(x)$, большее c_n , и в то же время все значения функции $f(x)$ не превосходят любого числа d_n , т. е. $f(x) \leq d_n$ при любом n . По лемме о вложенных стягивающихся отрезках существует точка M , принадлежащая всем отрезкам. Нетрудно видеть, что $M = \lim c_n = \lim d_n$. Переходя теперь к пределу в неравенстве $f(x) \leq d_n$, получим $f(x) \leq M$.

Докажем, что существует точка ξ отрезка $[a, b]$, в которой значение функции равно M : $f(\xi) = M$.

Действительно, как следует из свойства последовательности $\{c_n\}$, $c_n \rightarrow M$, существуют значения функции $f(x)$, сколь угодно мало отличающиеся от M . Допустим, что функция $f(x)$ ни в какой точке отрезка $[a, b]$ не принимает значение M . Тогда функция $\varphi(x) = \frac{1}{M - f(x)}$, непре-

рывная на отрезке $[a, b]$, будет неограниченной на этом отрезке, что противоречит теореме 1. Следовательно, существует такая точка ξ отрезка $[a, b]$, что $f(\xi) = M$. Таким образом, доказано, что среди значений непрерывной функции, принимаемых на отрезке $[a, b]$, имеется наибольшее. Докажем теперь, что среди значений непрерывной функции на отрезке $[a, b]$ имеется наименьшее. Действительно, рассмотрим функцию $-f(x)$. Она непрерывна на отрезке $[a, b]$ и, следовательно, на этом отрезке в некоторой точке η принимает свое наибольшее значение, обозначим его через $-m$. Отсюда следует, что функция $f(x)$ на отрезке $[a, b]$ в точке η принимает свое наименьшее значение, равное m .

Разность между наибольшим и наименьшим значениями функции на отрезке $[a, b]$ называется *колебанием непрерывной функции на этом отрезке*.

Прежде чем перейти к рассмотрению следующей теоремы, докажем лемму, на которую в дальнейшем неоднократно придется ссылаться.

Лемма. Если функция $f(x)$ непрерывна в точке x_0 и $f(x_0) \neq 0$, то существует окрестность точки x_0 , в которой функция сохраняет знак.

Доказательство. Допустим, что $f(x_0) > 0$. Тогда по определению непрерывности функции в точке x_0 для любого числа $\varepsilon > 0$ найдется число $\delta > 0$, такое, что неравенство $|f(x) - f(x_0)| < \varepsilon$ выполняется для всех x , удовлетворяющих условию $|x - x_0| < \delta$ или, что то же, неравенства

$$f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon \quad (1)$$

выполняются для всех точек x интервала $(x_0 - \delta, x_0 + \delta)$. Выбирая ε так, чтобы $0 < \varepsilon \leq \frac{f(x_0)}{2}$, и учитывая левое неравенство (1), получим $f(x) > \frac{f(x_0)}{2}$, откуда следует, что $f(x) > 0$ для всех точек x интервала $(x_0 - \delta, x_0 + \delta)$, что и требовалось доказать.

Если же $f(x_0) < 0$, то рассмотрим функцию $-f(x)$. Так как $-f(x) > 0$, то по доказанному существует окрестность точки x_0 , в которой $-f(x) > 0$ и, следовательно, $f(x) < 0$.

Теорема 3. Функция, непрерывная на отрезке $[a, b]$, принимает на этом отрезке все значения между любыми двумя ее значениями.

Рассмотрим значения функции в точках $x_1 = \alpha$ и $x_2 = \beta$ ($a \leq \alpha < \beta \leq b$), и пусть, например, $f(\alpha) < f(\beta)$. Возьмем какое-нибудь число μ между $f(\alpha)$ и $f(\beta)$.

Достаточно показать, что на отрезке $[\alpha, \beta]$ существует хотя бы одна точка ξ , в которой значение функции $f(x)$ равно μ : $f(\xi) = \mu$.

Докажем сначала предложение: если функция, непрерывная на отрезке $[a, b]$, на его концах имеет значения разных знаков, то на этом отрезке найдется точка ξ , в которой функция обращается в нуль.

Пусть, например, $f(a) < 0$, $f(b) > 0$.

Допустим, что предложение неверно, разделим отрезок $[a, b]$ пополам и выберем тот из полученных отрезков, на концах которого функция имеет значения разных знаков. Этот отрезок обозначим через $[a_1, b_1]$. Разделим отрезок $[a_1, b_1]$ пополам и выберем из двух отрезков тот, на концах которого функция $f(x)$ принимает значения разных знаков, обозначим его $[a_2, b_2]$ и т. д. Продолжая этот процесс неограниченно, получим последовательность $[a, b], [a_1, b_1], \dots, [a_n, b_n], \dots$ вложенных стягивающихся отрезков, причем на концах каждого отрезка функция имеет значения разных знаков. По лемме о вложенных стягивающихся отрезках существует точка ξ , принадлежащая всем отрезкам. Докажем, что $f(\xi) = 0$. В самом деле, если допустить, что $f(\xi) > 0$, то по доказанной лемме существует окрестность точки ξ , в которой $f(x) > 0$. В эту окрестность, при достаточно большом n , попадает отрезок $[a_n, b_n]$, в котором, следовательно, $f(x) > 0$, а это противоречит выбору последовательности отрезков. Аналогично доказывается, что $f(\xi)$ не может быть меньше нуля. Окончательно получаем, что $f(\xi) = 0$, причем точка ξ лежит на $[a, b]$, что противоречит допущению.

Таким образом, предложение доказано.

Докажем теперь теорему 3. Рассмотрим на $[\alpha, \beta]$ функцию $\varphi(x) = f(x) - \mu$. Тогда $\varphi(\alpha) = f(\alpha) - \mu < 0$, $\varphi(\beta) = f(\beta) - \mu > 0$. По доказанному предложению существует на $[\alpha, \beta]$ точка ξ , в которой $\varphi(\xi) = 0$. Но $\varphi(\xi) = f(\xi) - \mu$, следовательно, $f(\xi) = \mu$ и точка ξ лежит на отрезке $[\alpha, \beta]$.

Таких точек, где значение функции равно μ , может быть и несколько (см. на рис. 171 точки ξ_1, ξ_2, ξ_3).

С помощью этой теоремы и теоремы 2 легко показать, что если функция $f(x)$ непрерывна на отрезке $[a, b]$, то множество ее значений является от-

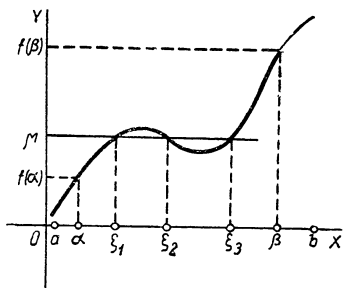


Рис. 171.

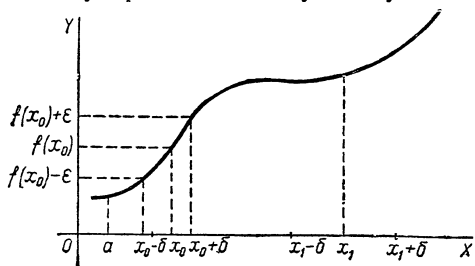


Рис. 172.

резком $[m, M]$, где m и M — соответственно наименьшее и наибольшее значения функции $f(x)$ на отрезке $[a, b]$.

Понятие о равномерной непрерывности функции. Пусть $f(x)$ — функция, непрерывная на промежутке, и x_0 — какая-то точка этого промежутка. Так как $f(x)$ непрерывна в точке x_0 , то для любого $\epsilon > 0$ найдется $\delta > 0$, такое, что $|f(x) - f(x_0)| < \epsilon$ при $|x - x_0| < \delta$ ¹⁾.

Ясно, что δ зависит от ϵ , но δ зависит также и от рассматриваемой точки x_0 . При выбранном $\epsilon > 0$ определяется $\delta > 0$, вообще говоря, различное для различных точек рассматриваемого промежутка. Если график функции $f(x)$ идет «круче» в окрестности рассматриваемой точки, то соответствующее этой точке δ будет меньше, чем δ , соответствующее точке, в окрестности которой график $f(x)$ более «пологий» (см. рис. 172).

¹⁾ Если точка x_0 совпадает с одним из концов промежутка, то рассматривается ее полукрестность.

Таким образом, при выбранном $\varepsilon > 0$ для каждой точки промежутка определяется соответствующее ей δ . Закрепим за каждой точкой свое положительное число δ . Если бы точек было конечное число, то из конечного числа положительных чисел δ можно было бы выбрать наименьшее число. Выбранное число δ было бы общим для всех этих точек в том смысле, что это δ , отвечающее выбранному ε , зависело бы только от ε и не зависело от x . Но из бесчисленного множества положительных чисел не всегда можно выбрать наименьшее число, как, например, среди членов последовательности $1, \frac{1}{2}, \dots, \frac{1}{n}, \dots$ наименьшего нет.

Если допустить, что из всех значений δ , соответствующих точкам рассматриваемого промежутка, можно выбрать наименьшее число $\delta, \delta > 0$, то тогда не будет необходимости закреплять «свое δ » за каждой точкой, наименьшее значение δ будет общим для всех точек этого промежутка, т. е. δ не будет зависеть от x . И тогда для любых двух точек x' и x'' рассматриваемого промежутка, удовлетворяющих условию $|x' - x''| < \delta$ (расстояние между которыми меньше δ), будет выполняться неравенство

$$|f(x') - f(x'')| < \varepsilon.$$

Предполагая, что такой выбор числа $\delta, \delta > 0$, общего для всех точек промежутка, возможен при любом $\varepsilon > 0$, то мы приходим к понятию равномерной непрерывности функции на промежутке.

Функция $f(x)$ называется *равномерно непрерывной на промежутке*, если для любого $\varepsilon > 0$ существует $\delta > 0$, такое, что для любых двух точек x' и x'' этого промежутка, удовлетворяющих условию $|x' - x''| < \delta$ (или, что то же, находящихся на расстоянии, меньшем δ), выполняется неравенство

$$|f(x') - f(x'')| < \varepsilon.$$

Теорема 4. Функция, непрерывная на отрезке, равномерно непрерывна на нем.

Доказательство. Сначала докажем, что если $f(x)$ непрерывна на отрезке $[a, b]$, то для любого $\varepsilon > 0$ можно разбить отрезок $[a, b]$ на конечное число неперекрывающихся отрезков¹⁾, на каждом из которых колебание функции будет меньше ε .

Доказательство от противного. Допустим, что существует число $\varepsilon > 0$, для которого такое разбиение отрезка $[a, b]$ невозможно. Тогда разделим отрезок $[a, b]$ пополам и выберем тот из полученных отрезков, для которого такое разбиение невозможно и обозначим его $[a_1, b_1]$. (Если бы для каждого из двух полученных отрезков было возможно такое разбиение, то оно было бы возможно и для отрезка $[a, b]$). Разделим отрезок $[a_1, b_1]$ пополам и выберем тот из полученных отрезков, для которого такое разбиение на частные отрезки невозможно, и т. д. Продолжая этот процесс, получим последовательность вложенных стягивающихся отрезков $[a, b], [a_1, b_1], [a_2, b_2], \dots, [a_n, b_n], \dots$, обладающих тем свойством, что ни один из них нельзя разбить на конечное число неперекрывающихся отрезков, так чтобы колебание функции на каждом частном отрезке было меньше ε . По лемме о вложенных стягивающихся отрезках существует точка ξ , общая всем отрезкам $[a_n, b_n]$. Так как функция $f(x)$ непрерывна в точке ξ , то для рассматриваемого ε ($\varepsilon > 0$) найдется $\delta > 0$, такое, что $|f(\xi) - f(x)| < \frac{\varepsilon}{2}$ для всех x из δ -окрестности точки ξ (или полуокрестности точки ξ , если ξ совпадает с одним из концов отрезка $[a, b]$). Тогда для любых двух точек x' и x'' рассматриваемой δ -окрестности ξ будет выполняться неравенство

$$\begin{aligned} |f(x') - f(x'')| &= |[f(x') - f(\xi)] + [f(\xi) - f(x'')]| \leq |f(x') - f(\xi)| + \\ &+ |f(\xi) - f(x'')| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

¹⁾ Два отрезка называются неперекрывающимися, если они или не имеют общих точек, или имеют только один общий конец.

т. е.

$$|f(x') - f(x'')| < \varepsilon.$$

В рассматриваемую δ -окрестность точки ξ при достаточно большом n , попадет отрезок $[a_n, b_n]$, и, следовательно, для любых двух точек x' и x'' этого отрезка будет справедливо неравенство $|f(x') - f(x'')| < \varepsilon$, т. е. колебание функции $f(x)$ на отрезке $[a_n, b_n]$ будет меньше ε , а это противоречит выбору последовательности отрезков.

Перейдем теперь непосредственно к доказательству теоремы, т. е. докажем, что функция $f(x)$, непрерывная на отрезке $[a, b]$, является равномерно непрерывной на этом отрезке.

Действительно, какое бы $\varepsilon > 0$ мы ни взяли, для него существует такое разбиение отрезка $[a, b]$ на конечное число отрезков, что колебание функции на каждом из этих отрезков меньше $\frac{\varepsilon}{2}$. Обозначим через δ наименьшую из длин отрезков деления и рассмотрим любые две точки x' и x'' отрезка $[a, b]$, находящиеся на расстоянии, меньшем δ : $|x' - x''| < \delta$. Либо эти точки принадлежат одному отрезку деления, либо принадлежат двум соседним отрезкам. В первом случае $|f(x') - f(x'')| < \frac{\varepsilon}{2} < \varepsilon$. Во втором случае, обозначая через x_0 общий конец соседних отрезков, получим:

$$\begin{aligned} |f(x') - f(x'')| &= \\ &= |[f(x') - f(x_0)] + [f(x_0) - f(x'')]| \leq \\ &\leq |f(x') - f(x_0)| + \\ &+ |f(x_0) - f(x'')| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Таким образом, для любого числа $\varepsilon > 0$ найдется такое число $\delta > 0$, что для любых двух точек x' и x'' отрезка $[a, b]$, расстояние между которыми меньше δ , выполняется неравенство $|f(x') - f(x'')| < \varepsilon$, что и требовалось доказать.

Следует иметь в виду, что теорема перестает быть верной для функций, непрерывных на интервалах или полуинтервалах.

Пример 1. $y = \sin \frac{1}{x}$. График этой функции изображен на рис. 168.

Функция $\sin \frac{1}{x}$ непрерывна на интервале $(0, a)$, где $a > 0$, но не является равномерно непрерывной на нем.

Действительно, какое бы мы малое $\delta > 0$ ни взяли, найдутся точки x' и x'' , достаточно близкие к 0, расстояние между которыми $|x' - x''| < \delta$, но $f(x') = 1$, а $f(x'') = -1$ и, следовательно, $|f(x') - f(x'')| = 2$. Поэтому для $\varepsilon < 2$ не существует соответствующего положительного числа δ .

Пример 2. $y = \frac{1}{x}$ (см. рис. 173).

Функция $\frac{1}{x}$ непрерывна на интервале $(0, 1)$, но не является равномерно непрерывной на нем. Это следует из того, что, какое бы мы число $\delta, \delta > 0$, ни взяли, существуют точки x', x'' , достаточно близкие к нулю, расстояние между которыми меньше δ , а модуль разности $|f(x') - f(x'')|$ больше любого наперед заданного числа.

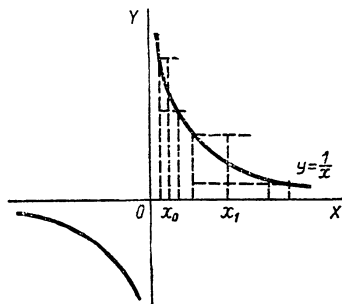


Рис. 173.

ГЛАВА IV

ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИИ ОДНОЙ ПЕРЕМЕННОЙ

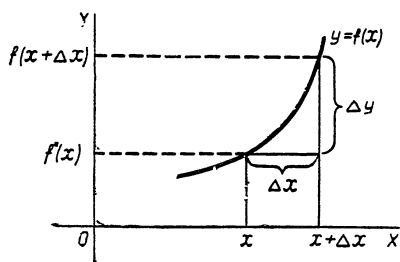
Раздел 1

ПРОИЗВОДНАЯ ФУНКЦИИ

§ 1. Приращение аргумента и функции

Рассмотрим функцию $y = f(x)$.

Пусть x — некоторое значение аргумента, $f(x)$ — соответствующее значение функции. От значения x переходим к другому значению аргумента x_1 . Разность $x - x_1$, обозначим ее Δx , называется *приращением аргумента*. Будем предполагать, что $\Delta x \neq 0$. Значение функции, соответствующее значению аргумента $x_1 = x + \Delta x$, равно $f(x + \Delta x)$. Разность $f(x + \Delta x) - f(x)$ называется *приращением функции* в точке x , соответствующим приращению аргумента Δx , и обозначается Δy или $\Delta f(x)$:



$$\Delta y = f(x + \Delta x) - f(x).$$

Рис. 174.

В зависимости от вида функции ее приращение Δy может быть равно нулю, положительно или отрицательно. Приращение аргумента также может быть различного знака. На рис. 174 Δx и Δy положительны.

Используя понятия «приращение аргумента» и «приращение функции», можно дать следующее определение непрерывности функции в точке. Функция $f(x)$ называется непрерывной в точке x , если бесконечно малому приращению аргумента соответствует бесконечно малое приращение функции.

§ 2. Задача о проведении касательной к кривой

Рассмотрим пару прямых с общей точкой M : неподвижную MT и подвижную MN (рис. 175). При этом неподвижная прямая MT называется предельным положением подвижной прямой MN , если угол φ между ними стремится к нулю. Пусть теперь M — некоторая точка кривой (рис. 176). Возьмем на этой кривой любую другую точку N и будем перемещать ее по кривой любым способом, неограниченно при-

ближая к точке M . (Точка N неограниченно приближается к точке M , если расстояние между ними стремится к нулю).

При этих условиях секущая MN , вообще говоря, меняет свое положение, вращаясь вокруг точки M . Если существует прямая, являющаяся предельным положением секущей MN при неограниченном приближении точки N по кривой к точке M , то эта прямая называется касательной к кривой в точке M .

Таким образом, касательной к кривой в ее точке M называется предельное положение секущей MN при неограничен-

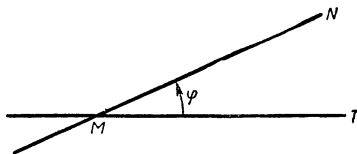


Рис. 175.

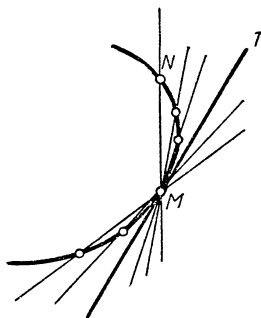


Рис. 176.

ном приближении точки N по кривой к точке M . Следует иметь в виду, что кривая в ее точке M может и не иметь касательной. Этот случай изображен на рис. 177.

Рассмотрим некоторую плоскую кривую с уравнением $y=f(x)$ и точку $M(x_0, y_0)$ этой кривой. Пусть кривая в точке M имеет неvertи-
кальную касательную (рис. 178). Напишем уравнение этой касательной.

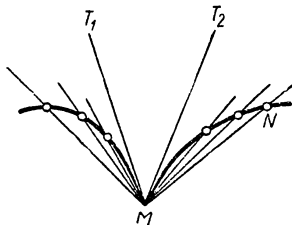


Рис. 177.

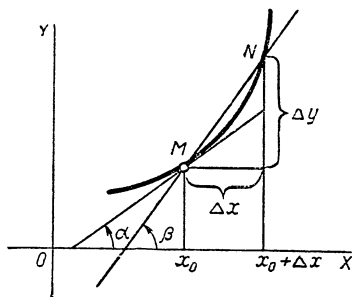


Рис. 178.

Значению аргумента $x_0 + \Delta x$ соответствует $f(x_0 + \Delta x) = y_0 + \Delta y$ и значит точка $N(x_0 + \Delta x, f(x_0 + \Delta x))$ кривой. Соединив точки M и N , получим секущую MN данной кривой. Уравнение секущей MN запишется в виде:

$$y - y_0 = \frac{\Delta y}{\Delta x} (x - x_0),$$

где $\Delta y = f(x_0 + \Delta x) - f(x_0)$ — приращение функции в точке x_0 .

Пусть теперь Δx стремится к нулю, тогда точка $N(x_0 + \Delta x, f(x_0 + \Delta x))$ по кривой стремится к точке $M(x_0, y_0)$, секущая MN , меняя свое положение, будет стремиться занять положение касательной к кривой в точке M .

Обозначим через β угол наклона к оси OX секущей MN , а через α — угол наклона касательной к кривой в точке M . Если $\Delta x \rightarrow 0$, то $\beta \rightarrow \alpha$ и $\operatorname{tg} \beta \rightarrow \operatorname{tg} \alpha$. Но $\operatorname{tg} \beta = \frac{\Delta y}{\Delta x}$, следовательно,

$$\operatorname{tg} \alpha = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}.$$

Таким образом, уравнение касательной к кривой в точке $M(x_0, y_0)$ имеет вид:

$$y - y_0 = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} (x - x_0).$$

И, следовательно, задача о проведении касательной к кривой в ее точке привела нас к вычислению предела отношения приращения функции к приращению аргумента при стремлении последнего к нулю.

§ 3. Определение производной функции в точке. Геометрический и механический смысл производной

Пусть функция $f(x)$ определена в некоторой окрестности точки x . Дадим аргументу x приращение Δx (при этом предполагается, что точка $x + \Delta x$ принадлежит области определения функции). Тогда функция получит приращение $\Delta y = f(x + \Delta x) - f(x)$. Производной функции $f(x)$ в точке x называется предел отношения приращения функции к приращению аргумента (т. е. $\frac{\Delta y}{\Delta x}$) при стремлении приращения аргумента к нулю, если этот предел существует и конечен.

Производная функции $y = f(x)$ в точке x обозначается символами

$$y', f'(x) \text{ или } \frac{dy}{dx}, \frac{df}{dx}^1).$$

Итак, по определению $y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ или, что то же,

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Как следует из определения, производная функции $f(x)$ в точке x есть число, зависящее от рассматриваемого значения x (но не зависящее от Δx). Рассматривая производную функции $f(x)$ в различных точках x , мы будем получать, вообще говоря, различные значения.

Таким образом, производная $f'(x)$ является функцией переменной x , определенной в области определения функции $f(x)$ или в части этой области²⁾.

Если фиксировать значение x , например положить $x = x_0$, то производная функции $f(x)$ в точке $x = x_0$ обозначается: $f'(x_0)$, $y'(x_0)$, $y'|_{x=x_0}$ или $\frac{df(x_0)}{dx}$.

Геометрический смысл производной. Рассматривая задачу о проведении касательной к кривой, мы получили, что $\operatorname{tg} \alpha = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$, т. е. $\operatorname{tg} \alpha = y'$.

¹⁾ Символы y' или $f'(x)$ (читаются: «игрек штрих» или «эф штрих от х») введены Лагранжем, а символы $\frac{dy}{dx}$ или $\frac{df}{dx}$ (читаются: «де игрек по де икс» или «де эф по де икс») введены Лейбницем.

Лагранж (1736—1813) — французский математик и механик, Лейбниц (1646—1718) — немецкий математик.

²⁾ Последнее, как увидим дальше, возможно, если функция $f(x)$ не во всех точках ее области определения имеет производную или, что то же, если не во всех точках области определения функции $f(x)$ существует конечный предел отношения $\frac{\Delta y}{\Delta x}$ при $\Delta x \rightarrow 0$.

Отсюда следует геометрический смысл производной функции $f(x)$. Производная функции $f(x)$ в точке x равна угловому коэффициенту касательной в точке $M(x, f(x))$ к кривой, заданной уравнением $y=f(x)$ [или, что то же, к графику функции $f(x)$].

В дальнейшем для сокращения формулировок вместо точки кривой будем называть только ее абсциссу, так как ордината этой точки однозначно определяется ее абсциссой.

Пример. Найдем производную функции $f(x)=x^2$ в точке x и выясним ее геометрический смысл.

Дадим x приращение Δx , тогда соответствующее ему приращение функции Δy запишется в виде:

$$\begin{aligned}\Delta y &= f(x + \Delta x) - f(x) = (x + \Delta x)^2 - x^2 = \\ &= 2x\Delta x + (\Delta x)^2.\end{aligned}$$

Откуда

$$\frac{\Delta y}{\Delta x} = 2x + \Delta x$$

и, следовательно, $y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = 2x$ или, что то же,

$$(x^2)' = 2x.$$

Таким образом, производная функции x^2 в любой точке x существует и равна $2x$, т. е. производная функции x^2 является функцией, определенной на всей числовой оси, равной $2x$.

В частности, при $x=1$ получим $f'(1)=2$.

Рассмотрим параболу $y=x^2$ (рис. 179). При $x=1$ $y=1$. Точка $M(1, 1)$ лежит на параболе. Так как $f'(1)=2$, то угловой коэффициент касательной к параболе в точке M равен двум и, следовательно, уравнение касательной к параболе в ее точке $M(1, 1)$ запишется в виде:

$$y - 1 = 2(x - 1),$$

т. е.

$$y = 2x - 1.$$

Положив $x=2$, найдем $f'(2)=4$, и, следовательно, угловой коэффициент касательной к параболе $y=x^2$ в точке $N(2, 4)$ равен 4.

Таким образом, $f'(x)=2x$ при любом x определяет угловой коэффициент касательной к параболе в соответствующей точке этой параболы.

В точке $O(0, 0)$ касательной к параболе является ось OX .

Заметим, что нельзя определить касательную как прямую, имеющую с кривой одну общую точку. В рассмотренном выше примере ось OY , имея с параболой одну общую точку, не является касательной к этой параболе. Определение касательной к кривой в данной точке было дано в § 2.

Под углом между двумя кривыми в точке их пересечения понимается угол между касательными к этим кривым в рассматриваемой точке.

Решим такую задачу. Определить, под каким углом синусоида наклонена к оси OX в начале координат (см. рис. 180).

Решение. Как следует из геометрического смысла производной, угловой коэффициент касательной к синусоиде в точке $x=0$ равен

$$\operatorname{tg} \alpha = y'|_{x=0}.$$

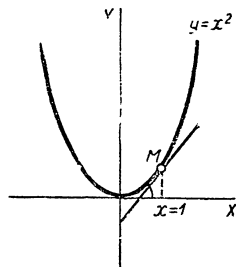


Рис. 179.

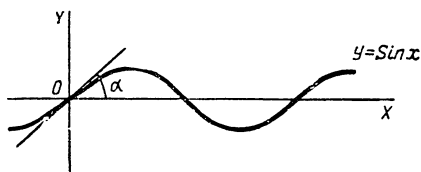


Рис. 180.

Сначала найдем $y' = (\sin x)'$ в любой точке x . Дадим x приращение Δx , тогда соответствующее приращение функции будет $\Delta y = \sin(x + \Delta x) - \sin x$.

$$\begin{aligned} y' &= \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\sin(x + \Delta x) - \sin x}{\Delta x} = \\ &= \lim_{\Delta x \rightarrow 0} \frac{2 \sin \frac{\Delta x}{2} \cos\left(x + \frac{\Delta x}{2}\right)}{\Delta x}. \end{aligned}$$

Или

$$y' = \lim_{\Delta x \rightarrow 0} \frac{\sin \frac{\Delta x}{2} \cos\left(x + \frac{\Delta x}{2}\right)}{\frac{\Delta x}{2}} = \lim_{\Delta x \rightarrow 0} \cos\left(x + \frac{\Delta x}{2}\right) = \cos x.$$

Окончательно получаем:

$$y' = (\sin x)' = \cos x.$$

Положив $x = 0$, найдем $y'|_{x=0} = \cos 0 = 1$.

Следовательно, $\operatorname{tg} \alpha = 1$, $\alpha = \frac{\pi}{4}$.

Заметим, что при нахождении производной непрерывной функции (ниже будет показано, что только для непрерывных функций можно ставить такую задачу) приходится раскрывать неопределенность вида $\frac{0}{0}$, так как $\Delta y \rightarrow 0$ при $\Delta x \rightarrow 0$.

Уравнения касательной и нормали к кривой. На основании ранее приведенных рассуждений получаем, что уравнение невертикальной касательной к кривой $y = f(x)$ в ее точке $M(x_0, y_0)$ можно записать в виде:

$$y - y_0 = f'(x_0)(x - x_0),$$

где $f'(x_0) = y'|_{x=x_0}$ — производная функции $f(x)$ в точке x_0 .

Нормалью к кривой в ее точке M называется прямая, проходящая через точку M перпендикулярно касательной к кривой в этой точке.

В силу условия перпендикулярности двух прямых можно записать уравнение нормали к кривой $y = f(x)$ в точке $M(x_0, y_0)$ в виде:

$$y - y_0 = -\frac{1}{f'(x_0)}(x - x_0),$$

если $f'(x_0) \neq 0$.

Пример. Написать уравнение касательной и нормали к синусоиде в точке $x = \frac{\pi}{4}$.

Решение. При $x = \frac{\pi}{4}$, $y = \sin \frac{\pi}{4} = \frac{\sqrt{2}}{2}$, следовательно, точка касания M имеет координаты $x = \frac{\pi}{4}$, $y = \frac{\sqrt{2}}{2}$. Так как в любой точке x , $(\sin x)' = \cos x$, то при $f'\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}$ и уравнение касательной к синусоиде в точке $M\left(\frac{\pi}{4}, \frac{\sqrt{2}}{2}\right)$ имеет следующий вид: $y - \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{2}\left(x - \frac{\pi}{4}\right)$.

Уравнение нормали к синусоиде в этой точке записывается в виде:

$$y - \frac{\sqrt{2}}{2} = -\sqrt{2}\left(x - \frac{\pi}{4}\right).$$

Односторонние производные функции в точке. Если рассматривать односторонние пределы отношения $\frac{\Delta y}{\Delta x}$ при $\Delta x \rightarrow 0$, то по аналогии с опре-

делением производной функции в точке x , можно сформулировать определение правой (левой) производной функции $f(x)$ в точке x .

Правой (левой) производной функции $f(x)$ в точке x называется предел отношения $\frac{\Delta y}{\Delta x}$ при $\Delta x \rightarrow 0$, $\Delta x > 0$ ($\Delta x < 0$), если этот предел существует и конечен.

Итак, по определению правая производная функции $f(x)$ в точке x есть число, равное $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x > 0}} \frac{\Delta y}{\Delta x}$; левая производная функции $f(x)$ в точке x есть число, равное $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x < 0}} \frac{\Delta y}{\Delta x}$.

Правая (левая) производная функции $f(x)$ в точке x совпадает с угловым коэффициентом правой (левой) касательной к графику функции $f(x)$ в точке x . Определить правую (левую) касательную к графику функции $f(x)$ в точке x предоставляем читателю.

Говорят, что функция $f(x)$ имеет *производную на интервале (a, b)* , если она имеет производную в каждой точке этого интервала. Функция $f(x)$ имеет *производную на отрезке $[a, b]$* , если она имеет производную на интервале (a, b) и, кроме того, имеет правую производную в точке a и левую производную в точке b .

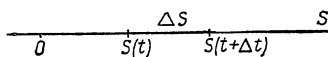


Рис. 181.

Механический смысл производной. Пусть некоторая материальная точка движется прямолинейно и задан закон ее движения $s = s(t)$, т. е. известно расстояние s точки от некоторого начала отсчета в каждый момент времени t .

Итак, в момент времени t пройденное расстояние равно $s(t)$ (см. рис. 181), в момент времени $t + \Delta t$ расстояние равно $s(t + \Delta t)$. За промежуток времени $(t, t + \Delta t)$ точка прошла путь $\Delta s = s(t + \Delta t) - s(t)$, $\frac{\Delta s}{\Delta t}$ — средняя скорость движения точки за промежуток времени $[t, t + \Delta t]$.

Средняя скорость движения на различных промежутках различна. Чем меньше промежуток времени Δt , тем точнее средняя скорость движения «характеризует» это движение в момент времени t .

Поэтому предел средней скорости движения при стремлении Δt к нулю называют *скоростью движения точки в данный момент времени t* и обозначают $v(t)$:

$$v(t) = \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t}.$$

Учитывая это, можно сказать, что производная $y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ характеризует скорость изменения функции в зависимости от изменения аргумента.

Необходимое условие существования производной.

Теорема. Если функция $f(x)$ имеет производную в точке x , то она непрерывна в этой точке.

Доказательство. По условию в точке x существует производная функции $f(x)$, т. е. существует конечный предел:

$$\lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x).$$

Из определения предела следует, что $\frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x) + \alpha$, где $\alpha \rightarrow 0$ при $\Delta x \rightarrow 0$, откуда $\Delta y = f(x + \Delta x) - f(x) = [f'(x) + \alpha] \Delta x$. И, следовательно, $\Delta y \rightarrow 0$ при $\Delta x \rightarrow 0$. Это и означает, что функция $f(x)$ непрерывна в точке x .

Таким образом, непрерывность функции в точке является необходимым условием существования производной функции в этой точке.

Заметим, что непрерывность функции в точке не является достаточным условием существования производной этой функции в рассматриваемой точке. Функция может быть непрерывной в точке, но не иметь в этой точке производной. Например, $y = |x|$ — непрерывная функция при $x = 0$ (график этой функции изображен на рис. 182). Дадим $x = 0$ приращение Δx . Функция $y = |x|$ получит приращение

$$\Delta y = |\Delta x|.$$

При этом $\frac{\Delta y}{\Delta x} = \begin{cases} 1, & \text{если } \Delta x > 0, \\ -1, & \text{если } \Delta x < 0. \end{cases}$ И, следовательно, предела отношения $\frac{\Delta y}{\Delta x}$ при $\Delta x \rightarrow 0$ не существует¹⁾. В точке $x = 0$ не существует касательной к графику функции $y = |x|$.

Заметим, что функция $y = |x|$ в точке $x = 0$ имеет правую и левую производные.

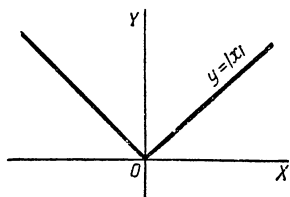


Рис. 182.

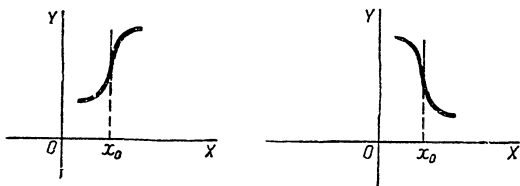


Рис. 183.

Действительно, $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x > 0}} \frac{\Delta y}{\Delta x} = 1$, т. е. правая производная в точке $x = 0$ равна 1; $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x < 0}} \frac{\Delta y}{\Delta x} = -1$, т. е. левая производная в этой точке равна -1 .

В точке $x = 0$ существуют правая и левая касательные к графику функции $|x|$. В этом случае точка $x = 0$ называется *угловой точкой* графика $|x|$.

Заметим еще, что в определении производной речь идет о конечном пределе отношения $\frac{\Delta y}{\Delta x}$ при $\Delta x \rightarrow 0$.

Если этот предел *бесконечен*, то говорят, что *производная не существует*. Геометрически мы имеем здесь случай *вертикальной касательной* (рис. 183).

Нахождение производной данной функции называется ее *дифференцированием*.

В этой главе рассматривается дифференциальное исчисление функции одной переменной, основная задача которого — вычисление производных, изучение их свойств и использование этих свойств для исследования функций.

§ 4. Производные некоторых элементарных функций

1. Пусть функция $y = f(x)$ на некотором интервале имеет постоянное значение, равное C :

$$y = C.$$

¹⁾ Здесь еще раз следует обратить внимание на то, что в определении производной приращение Δx стремится к нулю любым способом.

Тогда для любого x из рассматриваемого интервала $\Delta y = 0$, следовательно, $\frac{\Delta y}{\Delta x} = 0$, откуда получаем:

$$y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = 0, \\ (C)' = 0.$$

2. $y = x$. Пусть x — любое число ¹⁾. Дадим x приращение Δx , тогда приращение функции $\Delta y = \Delta x$ и $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = 1$, $(x)' = 1$.

3. $y = x^n$ (n — натуральное число).

Давая x приращение Δx , получим:

$$\Delta y = (x + \Delta x)^n - x^n = nx^{n-1}\Delta x + \frac{n(n-1)}{1 \cdot 2}x^{n-2}(\Delta x)^2 + \dots + (\Delta x)^n,$$

откуда

$$\frac{\Delta y}{\Delta x} = nx^{n-1} + \frac{n(n-1)}{1 \cdot 2}x^{n-2}\Delta x + \dots + (\Delta x)^{n-1}.$$

И, следовательно, $y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = nx^{n-1}$, т. е. $(x^n)' = nx^{n-1}$.

4. $y = \sin x$, $y' = (\sin x)' = \cos x$ (см. задачу § 3).

5. $y = \cos x$. Дадим x приращение Δx , соответствующее приращение y :

$$\Delta y = \cos(x + \Delta x) - \cos x,$$

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{-2 \sin \frac{\Delta x}{2} \sin \left(x + \frac{\Delta x}{2}\right)}{\Delta x} = \\ = - \lim_{\Delta x \rightarrow 0} \frac{\sin \frac{\Delta x}{2} \sin \left(x + \frac{\Delta x}{2}\right)}{\frac{\Delta x}{2}} = - \sin x, \\ (\cos x)' = - \sin x.$$

Решение примеров на нахождение производных значительно упрощается, если использовать общие правила дифференцирования, к рассмотрению которых мы переходим.

§ 5. Основные правила дифференцирования

Пусть функции $u = f(x)$ и $v = \varphi(x)$ имеют производные в точке x .

I. Производная алгебраической суммы функций равна алгебраической сумме производных этих функций, т. е. $(u \pm v)' = u' \pm v'$.

Действительно, пусть $y = u \pm v$. Дадим x приращение Δx , тогда функции u и v получат приращения, соответственно равные Δu и Δv , и функция y получит приращение Δy .

Новое значение функции u будет $u + \Delta u$, новое значение функции v будет $v + \Delta v$ и, следовательно,

$$\Delta y = [(u + \Delta u) \pm (v + \Delta v)] - (u \pm v) = \Delta u \pm \Delta v.$$

При этом

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} \pm \frac{\Delta v}{\Delta x}$$

и

$$y' = \lim_{\Delta x \rightarrow 0} \left(\frac{\Delta u}{\Delta x} \pm \frac{\Delta v}{\Delta x} \right) = \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} \pm \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x}$$

(так как предел каждого слагаемого существует и конечен).

Окончательно получим: $y' = (u \pm v)' = u' \pm v'$.

¹⁾ Условимся и в следующих примерах этого параграфа под x понимать любое число.

II. Производная произведения. Докажем, что $(uv)' = u'v + v'u$.

Дадим x приращение Δx , тогда функции u , v и $y = uv$ получают приращения. При этом

$$\Delta y = (u + \Delta u) \cdot (v + \Delta v) - uv = v\Delta u + u\Delta v + \Delta u \cdot \Delta v,$$

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \left(v \frac{\Delta u}{\Delta x} + u \frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta x} \Delta v \right).$$

Применяя теоремы о пределе суммы и произведения, будем иметь:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} v + \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} u + \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} \lim_{\Delta x \rightarrow 0} \Delta v.$$

Учитывая, что $\lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = u'$, $\lim_{\Delta x \rightarrow 0} \Delta v = 0$ (так как функция $v = \varphi(x)$ имеет производную в точке x , то она непрерывна в этой точке), получим:

$$y' = (uv)' = u'v + v'u,$$

что и требовалось доказать.

В частности, если $v = C$ ($C = \text{const}$), то

$$(Cu)' = (C)'u + Cu' = Cu',$$

т. е. постоянный множитель можно выносить за знак производной.

Пример 1. $y = 2x^3 + 5x^2 - 7$,

$$y' = (2x^3)' + (5x^2)' - (7)' = 2(x^3)' + 5(x^2)' - (7)' = 6x^2 + 10x.$$

Пример 2. $y = (7x^2 + 5x - 3) \cos x$,

$$y' = (7x^2 + 5x - 3)' \cos x + (7x^2 + 5x - 3) (\cos x)' = (14x + 5) \cos x - (7x^2 + 5x - 3) \sin x.$$

Формула дифференцирования произведения может быть обобщена на случай нескольких сомножителей.

Пусть u , v , w — функции, имеющие производные в точке x . Тогда $(uvw)' = [(uv)w]' = (uv)'w + (uv)w' = (u'v + uv')w + uvw' = u'vw + uv'w + uvw'$.

Окончательно получим:

$$(uvw)' = u'vw + uv'w + uvw'.$$

Аналогично для любого фиксированного числа сомножителей.

III. Производная частного. Рассмотрим функцию $y = \frac{u}{v}$. Предположим, кроме существования производных функций u и v в точке x , что функция $v = \varphi(x)$ отлична от нуля в рассматриваемой точке x . Докажем, что при этом

$$\left(\frac{u}{v}\right)' = \frac{u'v - v'u}{v^2}.$$

Действительно, приращение y , соответствующее приращению x , равно

$$\Delta y = \frac{u + \Delta u}{v + \Delta v} - \frac{u}{v} = \frac{v\Delta u - u\Delta v}{v(v + \Delta v)}$$

и

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{v \frac{\Delta u}{\Delta x} - u \frac{\Delta v}{\Delta x}}{v(v + \Delta v)}.$$

Применяя теоремы о пределах частного и произведения и учитывая непрерывность функции v в точке x ¹⁾, получим, что

$$y' = \left(\frac{u}{v}\right)' = \frac{u'v - v'u}{v^2}.$$

¹⁾ В силу непрерывности функции v в точке x $\Delta v \rightarrow 0$ при $\Delta x \rightarrow 0$.

В качестве примеров на дифференцирование частного рассмотрим вычисление производных тригонометрических функций.

$$1) (\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x} \right)' = \frac{(\sin x)' \cos x - (\cos x)' \sin x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}, \text{ т. е. } (\operatorname{tg} x)' = \sec^2 x.$$

$$2) (\operatorname{ctg} x)' = \left(\frac{\cos x}{\sin x} \right)' = \frac{(\cos x)' \sin x - (\sin x)' \cos x}{\sin^2 x} = -\frac{\sin^2 x + \cos^2 x}{\sin^2 x} = -\frac{1}{\sin^2 x}, \text{ т. е. } (\operatorname{ctg} x)' = -\operatorname{cosec}^2 x.$$

$$3) (\sec x)' = \left(\frac{1}{\cos x} \right)' = -\frac{(\cos x)'}{\cos^2 x} = \frac{\sin x}{\cos^2 x}, \text{ т. е. } (\sec x)' = \sec x \operatorname{tg} x.$$

$$4) (\operatorname{cosec} x)' = \left(\frac{1}{\sin x} \right)' = -\frac{(\sin x)'}{\sin^2 x} = -\frac{\cos x}{\sin^2 x},$$

$$(\operatorname{cosec} x)' = -\operatorname{cosec} x \operatorname{ctg} x^1).$$

IV. Производная сложной функции. Пусть $y = f(u)$, $u = \varphi(x)$, причем область изменения второй функции входит в область определения первой. Тогда y является сложной функцией независимой переменной x , u — промежуточная переменная. Предположим, что функция $\varphi(x)$ имеет производную в точке x , а функция $f(u)$, рассматриваемая как функция независимой переменной u , имеет производную в точке u , соответствующей точке x .

Докажем, что

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

или, что то же,

$$y' = f'(u) u', \quad (1)$$

т. е. докажем, что производная сложной функции по независимой переменной равна произведению производной этой функции по промежуточной переменной на производную промежуточной переменной по независимой переменной.

Доказательство. Дадим x приращение Δx , тогда u получит приращение Δu и y получит приращение Δy .

Запишем тождество

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x} \quad (2)$$

и найдем его предел при $\Delta x \rightarrow 0$. Если $\Delta x \rightarrow 0$, то $\Delta u \rightarrow 0$, так как $u = \varphi(x)$ — непрерывная функция в точке x . И, следовательно,

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x}$$

или $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$, что и требовалось доказать.

Следует иметь в виду, что приведенные выше рассуждения верны лишь в том случае, когда при достаточно малых Δx соответствующие приращения Δu отличны от нуля [иначе тождество (2) теряет смысл]. Можно доказать, что формула дифференцирования сложной функции остается верной и в этом случае, но мы это доказательство опускаем.

¹⁾ Нетрудно видеть, что для каждой из рассмотренных функций область определения самой функции и ее производной совпадают.

Формулу (1) иногда записывают и так:

$$y' = y'_u \cdot u'$$

(в записи производной y'_u значок « u » внизу подчеркивает, что дифференцирование y ведется по переменной u).

Рассмотрим примеры:

1) $y = (2 + x^2)^{10}$. Полагаем $y = u^{10}$, $u = 2 + x^2$, тогда $y' = 10u^9 u' = 10(2 + x^2)^9 \times (2 + x^2)' = 20x(2 + x^2)^9$.

2) $y = \sin x^2$. Полагая $y = \sin u$, $u = x^2$, получим $y' = (\sin u)' u' = \cos u (x^2)' = 2x \cos x^2$.

3) $y = \sin^3(2x + 3)$. Полагая $y = u^3$, $u = \sin v$, $v = 2x + 3$, получим $y' = (u^3)' u' = 3u^2 (\sin v)' = 3 \sin^2(2x + 3) \cos v \cdot v' = 3 \sin^2(2x + 3) \cos(2x + 3) \cdot 2 = 6 \sin^2(2x + 3) \cos(2x + 3)$.

При достаточном количестве упражнений необходимость в записи промежуточных переменных отпадает.

Дифференцирование сложной функции занимает центральное место в так называемой технике дифференцирования.

§ 6. Производная логарифмической функции

Логарифмическое дифференцирование. Рассмотрим функцию $y = \ln x$ и докажем, что для любого $x > 0$ $(\ln x)' = \frac{1}{x}$.

$$\begin{aligned} \text{Действительно, } \frac{\Delta y}{\Delta x} &= \frac{\ln(x + \Delta x) - \ln x}{\Delta x} = \ln\left(\frac{x + \Delta x}{x}\right)^{\frac{1}{\Delta x}} = \ln\left(1 + \frac{\Delta x}{x}\right)^{\frac{1}{\Delta x}} = \\ &= \ln\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x} \cdot \frac{1}{x}} = \frac{1}{x} \ln\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}. \end{aligned}$$

$$\text{При этом } y' = \lim_{\Delta x \rightarrow 0} \frac{1}{x} \ln\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}} = \frac{1}{x} \lim_{\Delta x \rightarrow 0} \ln\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}.$$

Из соотношения $\lim_{\Delta x \rightarrow 0} \left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}} = e$, учитывая непрерывность логарифмической функции, получаем:

$$\lim_{\Delta x \rightarrow 0} \ln\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}} = \ln \lim_{\Delta x \rightarrow 0} \left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}} = \ln e = 1$$

и, следовательно, $y' = (\ln x)' = \frac{1}{x}$.

Если же $y = \log_a x$ ($a > 0$, $a \neq 1$), то учитывая равенство $\log_a x = \frac{\ln x}{\ln a}$, получим: $(\log_a x)' = \frac{1}{x \ln a}$.

Найдем теперь производную функции $y = \ln|x|$. Эта функция определена при всех x , кроме $x = 0$. График ее изображен на рис. 184.

$$y = \ln|x| = \begin{cases} \ln x & \text{при } x > 0, \\ \ln(-x) & \text{при } x < 0. \end{cases}$$

Так как $(\ln x)' = \frac{1}{x}$ и $[\ln(-x)]' = \frac{(-x)'}{-x} = \frac{1}{x}$ (последнее равенство записано на основании дифференцирования сложной функции, здесь $u = -x$), то окончательно получаем:

$$(\ln|x|)' = \frac{1}{x}.$$

Учитывая полученное равенство, найдем производную функции $y = \ln |f(x)|$ или, что то же, функции $y = \ln |u|$, где $u = f(x)$. При этом предполагается, что функция $f(x)$ не обращается в нуль в рассматриваемой точке x и имеет производную в этой точке.

Дифференцируя сложную функцию, получим:

$$y' = (\ln |u|)' u' = \frac{1}{u} \cdot u' = \frac{f'(x)}{f(x)},$$

т. е.

$$(\ln |f(x)|)' = \frac{f'(x)}{f(x)}. \quad (1)$$

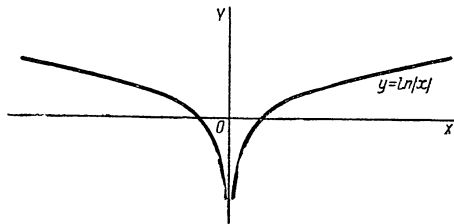


Рис. 184.

Отношение $\frac{f'(x)}{f(x)}$ называется *логарифмической производной* функции $f(x)$. Как это следует из равенства (1), логарифмическая производная функции равна производной логарифма модуля этой функции.

Способ логарифмического дифференцирования состоит в том, что сначала находят логарифмическую производную функции, а затем производную самой функции по формуле:

$$f'(x) = (\ln |f(x)|)' f(x). \quad (2)$$

Для упрощения записи условимся при логарифмическом дифференцировании знак модуля у $f(x)$ опускать.

Рассмотрим пример.

$$y = \frac{(x-1)^{10} e^{\sin x}}{\sqrt{(1+x^2)^8}}.$$

Логарифмируя, запишем:

$$\ln y = 10 \ln (x-1) + \sin x - \frac{3}{2} \ln (1+x^2).$$

Дифференцируя, получим:

$$\frac{y'}{y} = \frac{10}{x-1} + \cos x - \frac{3}{2} \cdot \frac{2x}{1+x^2},$$

откуда

$$y' = \left(\frac{10}{x-1} + \cos x - \frac{3x}{1+x^2} \right) y = \left(\frac{10}{x-1} + \cos x - \frac{3x}{1+x^2} \right) \frac{(x-1)^{10} e^{\sin x}}{\sqrt{(1+x^2)^8}}.$$

(Метод логарифмического дифференцирования в данном случае применим, если $x \neq 1$. Можно показать, что производная рассматриваемой функции существует при любом x , т. е. в области определения самой функции.)

§ 7. Производные степенных и показательных функций

1) $y = x^\alpha$ (α — любое действительное число).

При любом x , отличном от нуля,

$$\begin{aligned} \ln y &= \alpha \ln x, \\ \frac{y'}{y} &= \alpha \cdot \frac{1}{x}, \quad y' = \frac{\alpha}{x} y = \alpha x^{\alpha-1}, \\ (x^\alpha)' &= \alpha x^{\alpha-1}. \end{aligned}$$

¹⁾ Если $\alpha > 1$, то можно непосредственно из определения производной получить, что производная функции x^α при $x=0$ равна нулю.

В частности, при $\alpha = -1$ имеем:

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2},$$

при $\alpha = \frac{1}{2}$ получим:

$$(\sqrt{x})' = \frac{1}{2\sqrt{x}}.$$

Пример. $y = \sqrt{2x^2 + 3x + 7}$,

$$y' = \frac{(2x^2 + 3x + 7)'}{2\sqrt{2x^2 + 3x + 7}} = \frac{4x + 3}{2\sqrt{2x^2 + 3x + 7}}.$$

2) $y = a^x$ ($a > 0$, $a \neq 1$).

При любом x

$$\ln y = x \ln a,$$

$$\frac{y'}{y} = \ln a; \quad y' = \ln a \cdot y = a^x \ln a,$$

$$(a^x)' = a^x \ln a.$$

Положив $a = e$, получим $(e^x)' = e^x$.

Пример 1. $y = e^{\sin^2 x}$, $y' = e^{\sin^2 x} (\sin^2 x)' = 2 \sin x \cos x e^{\sin^2 x} = \sin 2x e^{\sin^2 x}$.

Пример 2. $y = 2^{\ln x}$, $y' = 2^{\ln x} (\ln x)' \ln 2 = 2^{\ln x} \frac{\ln 2}{x}$.

3) **Производная показательно-степенной функции.** Пусть $u = f(x)$ и $v = \varphi(x)$ — функции, имеющие производные в точке x , $f'(x) > 0$. Найдем производную функции $y = u^v$. Логарифмируя, получим:

$$\ln y = v \ln u,$$

$$\frac{y'}{y} = v' \ln u + v \frac{u'}{u},$$

откуда

$$y' = u^v \left(v \frac{u'}{u} + v' \ln u \right)$$

или

$$(u^v)' = v u^{v-1} u' + u^v v' \ln u.$$

§ 8. Производные обратных функций

Производные обратных тригонометрических функций. Сначала рассмотрим производные обратных тригонометрических функций.

Пусть $y = \arcsin x$ ($-1 < x < 1$). Тогда $x = \sin y$. Дифференцируя по x ($-1 < x < 1$), получим:

$$1 = \cos y \cdot y',$$

откуда $y' = \frac{1}{\cos y}$, причем y принадлежит интервалу $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Так как на этом интервале

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2} \quad (-1 < x < 1),$$

то, записав вместо $\cos y$ его выражение через x , окончательно получим:

$$(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}} \quad (-1 < x < 1).$$

Пусть $y = \arccos x$ ($-1 < x < 1$), тогда $x = \cos y$ и дифференцируя по x ($-1 < x < 1$), получим

$$1 = -\sin y \cdot y'.$$

Так как для всех y интервала $(0, \pi)$ имеем $\sin y > 0$, то

$$y' = -\frac{1}{\sin y} = -\frac{1}{\sqrt{1-\cos^2 y}} = -\frac{1}{\sqrt{1-x^2}} \quad (-1 < x < 1).$$

Пусть теперь $y = \operatorname{arctg} x$ $(-\infty < x < +\infty)$.

Тогда в результате дифференцирования по x , $-\infty < x < +\infty$, функции $x = \operatorname{tg} y$ получим:

$$1 = \sec^2 y \cdot y',$$

откуда

$$y' = \frac{1}{\sec^2 y} = \frac{1}{1 + \operatorname{tg}^2 y} = \frac{1}{1 + x^2},$$

т. е.

$$(\operatorname{arctg} x)' = \frac{1}{1 + x^2} \quad (-\infty < x < +\infty).$$

Получить равенство

$$(\operatorname{arcsctg} x)' = -\frac{1}{1 + x^2} \quad (-\infty < x < +\infty)$$

предоставляем читателю.

Пусть теперь требуется найти производную функции $y = f(x)$ при условии, что обратная ей функция $x = \varphi(y)$ имеет производную, отличную от нуля, в соответствующей точке.

Для решения этой задачи дифференцируем функцию $x = \varphi(y)$ по x . Тогда имеем: $1 = \varphi'(y) y'$. Так как $\varphi'(y) \neq 0$ в рассматриваемой точке, то

$$y' = \frac{1}{\varphi'(y)}$$

или

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}},$$

т. е. производная функции, обратной данной, обратна по величине производной данной функции.

Используя этот результат, можно найти, например, производные обратных тригонометрических функций. Пусть $y = \operatorname{arcsin} x$, тогда $x = \sin y$ и $\frac{dy}{dx} = \frac{1}{\cos y}$ при y из интервала $(-\frac{\pi}{2}, \frac{\pi}{2})$.

Выражая, наконец, $\cos y$ через x , как это уже было сделано, получим:

$$(\operatorname{arcsin} x)' = \frac{1}{\sqrt{1-x^2}} \quad (-1 < x < 1).$$

Рассуждая аналогично, найдем производные $\operatorname{arccos} x$ и $\operatorname{arctg} x$.

Пример 1. $y = \operatorname{arctg} e^x$.

$$y' = \frac{(e^x)'}{1 + e^{2x}} = \frac{e^x}{1 + e^{2x}}.$$

Пример 2. $y = \operatorname{arcsin} (1 + \ln x)$.

$$y' = \frac{(1 + \ln x)'}{\sqrt{1 - (1 + \ln x)^2}} = \frac{1}{x \sqrt{1 - (1 + \ln x)^2}}.$$

Заканчивая рассмотрение формул дифференцирования, отметим, что, пользуясь полученными формулами производных основных элементарных функций и правилами дифференцирования, можно найти производную любой элементарной функции, т. е. функции, которая получается из основных элементарных функций с помощью конечного числа алгебраических действий и применением конечного числа операций вычисления функции от функции.

§ 9. Таблица формул дифференцирования

1. $y = C$ ($C = \text{const}$)	$y' = 0$
2. $y = x$	$y' = 1$
3. $y = x^a$	$y' = ax^{a-1}$
$y = \frac{1}{x}$	$y' = -\frac{1}{x^2}$
$y = \sqrt{x}$	$y' = \frac{1}{2\sqrt{x}}$
4. $y = a^x$ ($a > 0, a \neq 1$)	$y' = a^x \ln a$
$y = e^x$	$y' = e^x$
5. $y = \log_a x$ ($a > 0, a \neq 1$)	$y' = \frac{1}{x \ln a}$
$y = \ln x$	$y' = \frac{1}{x}$
6. $y = \sin x$	$y' = \cos x$
7. $y = \cos x$	$y' = -\sin x$
8. $y = \operatorname{tg} x$	$y' = \sec^2 x$
9. $y = \operatorname{ctg} x$	$y' = -\operatorname{cosec}^2 x$
10. $y = \sec x$	$y' = \sec x \operatorname{tg} x$
11. $y = \operatorname{cosec} x$	$y' = -\operatorname{cosec} x \operatorname{ctg} x$
12. $y = \arcsin x$	$y' = \frac{1}{\sqrt{1-x^2}}$
13. $y = \arccos x$	$y' = -\frac{1}{\sqrt{1-x^2}}$
14. $y = \operatorname{arctg} x$	$y' = \frac{1}{1+x^2}$
15. $y = \operatorname{arcctg} x$	$y' = -\frac{1}{1+x^2}$

§ 10. Замечания о дифференцировании неявных функций

Общее выражение производных неявных функций будет дано в гл. 7. Но уже теперь, используя дифференцирование сложной функции, для любой конкретно заданной неявной функции можно найти производную. (При этом предполагается, что производная этой функции существует). В качестве примера найдем производную y' функции y от x , неявно заданной уравнением

$$x + \sqrt{xy} + y = a.$$

Решение. Дифференцируя $x + \sqrt{xy} + y = a$ по x , получим соотношение

$$1 + \frac{y + xy'}{2\sqrt{xy}} + y' = 0,$$

из которого определим y' :

$$\begin{aligned} y + 2\sqrt{xy} + y'(x + 2\sqrt{xy}) &= 0, \\ y' &= -\frac{y + 2\sqrt{xy}}{x + 2\sqrt{xy}}. \end{aligned}$$

Выражая $\sqrt{xy}$ из уравнения через x и y :

$$\sqrt{xy} = a - x - y,$$

получим:

$$y' = -\frac{2a - 2x - y}{2a - 2y - x}.$$

Задача. Написать уравнение касательной к эллипсу $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ в его точке $M(x_0, y_0)$.

Как известно, уравнение касательной к кривой $y = f(x)$ в ее точке $M(x_0, y_0)$ записывается в виде:

$$y - y_0 = f'(x_0)(x - x_0).$$

Уравнение эллипса можно рассматривать как неявное задание y как функции x . Дифференцируя $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ по x , получим:

$$\frac{2x}{a^2} + \frac{2y}{b^2} y' = 0,$$

откуда

$$y' = -\frac{b^2 x}{a^2 y}.$$

Так как при $x = x_0$ имеем по условию $y = y_0$, то

$$y' = f'(x_0) = -\frac{b^2 x_0}{a^2 y_0}.$$

Подставляя значение производной y' в уравнение касательной, получим:

$$y - y_0 = -\frac{b^2 x_0}{a^2 y_0} (x - x_0)$$

или, что то же,

$$b^2 x_0 x + a^2 y_0 y = b^2 x_0^2 + a^2 y_0^2,$$

откуда, после деления этого уравнения на произведение $a^2 b^2$, получим:

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = \frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1$$

(последнее равенство записано с учетом того, что точка $M(x_0, y_0)$ лежит на эллипсе).

Окончательно уравнение касательной к эллипсу запишется в виде:

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1.$$

Раздел 2

ДИФФЕРЕНЦИАЛ. ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИАЛЫ ВЫСШИХ ПОРЯДКОВ

§ 11. Дифференциал функции

Пусть функция $y = f(x)$ имеет производную в точке x ,

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(x).$$

Тогда $\frac{\Delta y}{\Delta x} = f'(x) + \alpha$, где $\alpha \rightarrow 0$ при $\Delta x \rightarrow 0$, и, следовательно, $\Delta y = f'(x) \Delta x + \alpha \Delta x$.

Приращение функции записано в виде суммы двух слагаемых. Первое слагаемое $f'(x) \Delta x$ является при $\Delta x \rightarrow 0$ бесконечно малой первого порядка с Δx (при $f'(x) \neq 0$). Это слагаемое линейно относительно Δx ; второе слагаемое $\alpha \Delta x$ при $\Delta x \rightarrow 0$ — бесконечно малая более высокого порядка, чем Δx ($\frac{\alpha \Delta x}{\Delta x} = \alpha \rightarrow 0$ при $\Delta x \rightarrow 0$).

Таким образом, первое слагаемое является главной частью приращения функции.

Функция $y = f(x)$ называется *дифференцируемой в точке x* , если приращение функции в точке x можно записать как сумму главной линейной относительно Δx части приращения и бесконечно малой более высокого порядка, чем Δx , т. е. если $\Delta y = A \Delta x + \alpha \Delta x$, где A — число, (A не зависит от Δx), $\alpha \rightarrow 0$ при $\Delta x \rightarrow 0$.

Из приведенных выше рассуждений следует: если функция $f(x)$ имеет производную в точке x , то она дифференцируема в этой точке, причем $A = f'(x)$.

Справедливо и обратное утверждение: если $f(x)$ дифференцируема в точке x , то она имеет производную в этой точке.

Действительно, пусть $\Delta y = A\Delta x + \alpha$, где A не зависит от Δx и $\alpha \rightarrow 0$ при $\Delta x \rightarrow 0$.

Тогда $\frac{\Delta y}{\Delta x} = A + \frac{\alpha}{\Delta x}$, а значит, $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = A$, т. е. $f'(x) = A$.

Итак, понятия «функция, дифференцируемая в точке» и «функция, имеющая производную в точке» — равносильны.

Функция $f(x)$ называется *дифференцируемой на интервале*, если она дифференцируема в каждой точке этого интервала или, что то же, если она имеет производную в каждой точке интервала.

Пусть функция $f(x)$ дифференцируема в точке x .

Дифференциалом функции $f(x)$ в точке x называется главная, линейная относительно Δx , часть приращения функции. Дифференциал функции $y = f(x)$ обозначается через dy или $df(x)$. Из определения следует, что

$$dy = f'(x) \Delta x. \quad (1)$$

Разность $\Delta y - dy = \alpha \Delta x$ — бесконечно малая более высокого порядка, чем Δx .

Если $y = x$, то $dy = dx = (x)' \Delta x = \Delta x$ (дифференциал независимой переменной равен приращению этой переменной). Учитывая это, запишем:

$$dy = f'(x) dx. \quad (2)$$

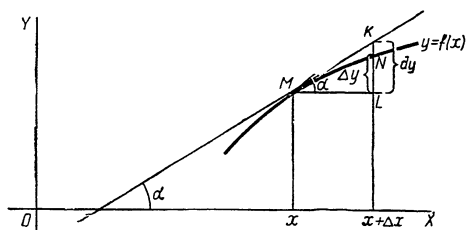


Рис. 185.

Следовательно, дифференциал функции $f(x)$ в точке x равен произведению производной функции в этой точке на дифференциал независимой переменной.

Из соотношения (2) производную $f'(x)$ можно определить как отношение дифференциалов: $f'(x) = \frac{dy}{dx}$. Таким образом, символ

производной $\frac{dy}{dx}$ можно рассматривать как отношение дифференциалов¹⁾.

Заметим, что dy зависит от x и Δx , $y' = f'(x)$ зависит только от x .

Геометрический смысл дифференциала. Рассмотрим график функции $y = f(x)$ (рис. 185). Из треугольника MLK имеем $KL = dy = \operatorname{tg} \alpha \cdot \Delta x$ или $dy = y' \Delta x$.

Таким образом, дифференциал функции $f(x)$ в точке x равен приращению ординаты касательной к графику этой функции в рассматриваемой точке.

Свойства дифференциала. Пусть $u = f(x)$, $v = \varphi(x)$ — функции, дифференцируемые в точке x . Непосредственно из определения дифференциала и свойств производных вытекают следующие свойства дифференциала.

$$1) d(u \pm v) = (u \pm v)' dx = u' dx \pm v' dx = du \pm dv,$$

$$d(u \pm v) = du \pm dv.$$

$$2) d(uv) = (uv)' dx = (u'v + uv') dx = vdu + u dv,$$

$$d(uv) = vdu + u dv.$$

¹⁾ Принятая ранее для обозначения производной символическая дробь $\frac{dy}{dx}$ после введения понятия дифференциала становится реальной дробью — отношением дифференциалов.

В частности, если $v=C$ ($C=\text{const}$), то $d(Cu)=Cdu$.

$$3) d\left(\frac{u}{v}\right)=\left(\frac{u}{v}\right)' dx=\frac{u'v-v'u}{v^2} dx=\frac{vdu-udv}{v^2},$$

$$d\left(\frac{u}{v}\right)=\frac{vdu-udv}{v^2}.$$

4) **Дифференциал сложной функции.** Инвариантная форма записи дифференциала. Пусть $y=f(x)$, $x=\varphi(t)$, т. е. y является сложной функцией независимой переменной t ; x — промежуточная переменная. Предположим, что функция $\varphi(t)$ дифференцируема в точке t , а функция $f(x)$, рассматриваемая как функция независимой переменной x , дифференцируема в точке x , соответствующей точке t . Как известно, при этих условиях сложная функция y аргумента t имеет производную в точке t , т. е. является дифференцируемой в точке t . Ее дифференциал в этой точке будет иметь вид: $dy=y'_t dt$ или, так как $y'_t=f'(x)\varphi'(t)$, то $dy=f'(x)\varphi'(t)dt=f'(x)dx$, т. е.

$$dy=f'(x)dx. \quad (3)$$

Сравнивая равенства (2) и (3), замечаем, что форма (2) записи дифференциала не зависит от того, будет ли x независимой переменной или функцией другой переменной, в связи с чем эта форма записи называется *инвариантной формой записи дифференциала*.

Следует иметь в виду, что если x — независимая переменная, то $dx=\Delta x$, если же x зависит от другой переменной, то $dx\neq\Delta x$, т. е. форма (1) записи дифференциала инвариантной не является.

Из равенства (3) снова получаем выражение производной в виде отношения дифференциалов: $f'(x)=\frac{dy}{dx}$. Эта запись производной в виде отношения дифференциалов справедлива не только, когда x независимая переменная, как было получено выше, но также и в том случае, когда x зависит от другой переменной. Выражение производной через отношение дифференциалов в дальнейшем неоднократно будет использоваться.

Заметим, что из определения и свойств дифференциала следует также, что нахождение производных или дифференциалов по существу представляет собой одну и ту же задачу, поэтому она и называется *дифференцированием*.

Приближенные вычисления с помощью дифференциала. Из определения дифференциала функции $y=f(x)$ в точке x следует, что он зависит линейно от Δx и является главной частью приращения функции Δy . Само же приращение функции Δy зависит от Δx обычно более сложным образом. Например, если $y=x^3$, то $\Delta y=(x+\Delta x)^3-x^3=3x^2\Delta x+3x(\Delta x)^2+(\Delta x)^3$, в то время как $dy=3x^2\Delta x$.

Кроме того, для вычисления дифференциала можно воспользоваться равенством $dy=f'(x)dx$. Поэтому во многих задачах вместо приращения функции в точке рассматривают дифференциал функции в этой точке, если $f'(x)\neq 0$, полагая $\Delta y\approx dy$. При этом *абсолютная погрешность* dy равна $|\Delta y-dy|$. Она является бесконечно малой более высокого порядка, чем Δx .

Относительной погрешностью dy называется $\left|\frac{\Delta y-dy}{dy}\right|$. Покажем, что относительная погрешность при условии, что $f'(x)\neq 0$, также бесконечно малая при $\Delta x\rightarrow 0$.

$$\text{В самом деле, } \left|\frac{\Delta y-dy}{dy}\right|=\left|\frac{\Delta\Delta x}{f'(x)\Delta x}\right|\rightarrow 0 \text{ при } \Delta x\rightarrow 0.$$

Таким образом, если заменить Δy величиной dy , то не только абсолютная погрешность стремится к нулю при $\Delta x\rightarrow 0$, но также и относительная погрешность стремится к нулю.

Перейдем теперь к рассмотрению примеров. Покажем, что если α очень мало, то можно использовать следующие приближенные формулы:

$$\sqrt{1+\alpha} \approx 1 + \frac{\alpha}{2},$$

$$\frac{1}{\sqrt{1-\alpha}} \approx 1 - \frac{\alpha}{2},$$

$$\ln(1+\alpha) \approx \alpha.$$

Действительно, возьмем $y = \sqrt{x}$. Тогда

$$\Delta y = \sqrt{x + \Delta x} - \sqrt{x} \approx dy,$$

$$\sqrt{x + \Delta x} - \sqrt{x} \approx (\sqrt{x})' \Delta x$$

или

$$\sqrt{x + \Delta x} - \sqrt{x} \approx \frac{1}{2\sqrt{x}} \Delta x,$$

откуда, положив $x = 1$, $\Delta x = \alpha$, получим:

$$\sqrt{1+\alpha} \approx 1 + \frac{\alpha}{2}.$$

В частности, при $\alpha = 0,0003$, найдем:

$$\sqrt{1,0003} \approx 1,00015.$$

Аналогично, рассматривая $y = \frac{1}{\sqrt{x}}$ и $y = \ln x$, получим две другие приближенные формулы.

§ 12. Производные и дифференциалы высших порядков

Пусть $y = f(x)$ — дифференцируемая функция на некотором интервале, тогда, дифференцируя ее, получим функцию $y' = f'(x)$ — производную (первую производную). Допустим, что эту функцию можно дифференцировать в точке x рассматриваемого интервала. Производная функции $f'(x)$ в точке x называется *второй производной функции $f(x)$* , или *производной второго порядка* в этой точке. Она обозначается: y'' , $f''(x)$ или $\frac{d^2y}{dx^2}$, $\frac{d^2f}{dx^2}$.

Таким образом, по определению, $y'' = (y')'$ или

$$\frac{d^2y}{dx^2} = \frac{d}{dx} \left(\frac{dy}{dx} \right).$$

Если функция $f'(x)$ дифференцируема на некотором интервале, то на этом интервале определяется функция $f''(x)$.

Производная от второй производной, если она существует, называется *третьей производной* и т. д., n -й производной, или *производной n -го порядка функции $f(x)$* в точке x (на некотором интервале) называется производная от производной $(n-1)$ -го порядка в этой точке (на этом интервале), она обозначается: $y^{(n)}$, $f^{(n)}(x)$ или $\frac{d^ny}{dx^n}$, $\frac{d^nf(x)}{dx^n}$.

Итак, по определению $y^{(n)} = (y^{(n-1)})'$ или, что то же,

$$\frac{d^ny}{dx^n} = \frac{d}{dx} \left(\frac{d^{n-1}y}{dx^{n-1}} \right).$$

Заметим, что при $n = 1, 2, 3$ порядок производной обозначается штрихами, при $n \geq 4$ обозначается арабской или римской цифрами, записанными в скобках. Читается: y''' — «игрек три штриха» или $\frac{d^3y}{dx^3}$ — «де три игрек по де икс куб» и т. д.

Пример: $y = x^4 + 5x + 4$, $y' = 4x^3 + 5$, $y'' = 12x^2$, $y''' = 24x$, $y^{(IV)} = 24$, $y^{(V)} = y^{(VI)} = \dots = 0$.

¹⁾ При некоторых условиях на функцию $y = f(x)$.

Вторая производная имеет простой *механический смысл*. Если задан закон прямолинейного движения $s=f(t)$, то, как известно, $\frac{ds}{dt}=v(t)$ — скорость в момент времени t .

$$\frac{d^2s}{dt^2} = \frac{d}{dt} \left(\frac{ds}{dt} \right) = \frac{dv}{dt},$$

но $\frac{\Delta v}{\Delta t}$ — среднее ускорение за промежуток времени $(t, t + \Delta t)$, $\frac{\Delta v}{\Delta t} = \omega_{\text{ср}}$. Предел среднего ускорения при $\Delta t \rightarrow 0$ называется *ускорением в момент времени t* : $\lim_{\Delta t \rightarrow 0} \frac{\Delta v}{\Delta t} = \frac{dv}{dt} = \omega(t)$.

Следовательно, $\frac{d^2s}{dt^2} = \omega(t)$ — вторая производная пути по времени, есть ускорение движения в момент времени t .

Примеры нахождения производных высших порядков.

1) $y = x^\alpha$, α — любое действительное число,

$$y' = \alpha x^{\alpha-1}, y'' = \alpha(\alpha-1)x^{\alpha-2}, \dots, y^{(n)} = \alpha(\alpha-1) \dots [\alpha-(n-1)] x^{\alpha-n}.$$

2) $y = a^x$, $a > 0$, $a \neq 1$,

$$y' = a^x \ln a, y'' = a^x (\ln a)^2, \dots, y^{(n)} = a^x (\ln a)^n.$$

3) $y = \sin x$,

$$y' = \cos x = \sin \left(x + \frac{\pi}{2} \right),$$

$$y'' = \sin \left(x + \frac{\pi}{2} + \frac{\pi}{2} \right) \left(x + \frac{\pi}{2} \right)' = \sin \left(x + \frac{2\pi}{2} \right),$$

$$y''' = \sin \left(x + \frac{2\pi}{2} + \frac{\pi}{2} \right) \left(x + \frac{2\pi}{2} \right)' = \sin \left(x + \frac{3\pi}{2} \right),$$

.....

И вообще

$$y^{(n)} = (\sin x)^{(n)} = \sin \left(x + \frac{n\pi}{2} \right)$$

(при этом заданную функцию считают производной нулевого порядка).

4) $y = \cos x$.

Замечая, что $(\cos x)' = -\sin x = \cos \left(x + \frac{\pi}{2} \right)$, аналогично получаем: $y^{(n)} = \cos \left(x + n \frac{\pi}{2} \right)^{1)}$.

Общие правила нахождения высших производных. Формула Лейбница. Пусть $u=f(x)$ и $v=\varphi(x)$ — функции, имеющие производные n -го порядка на некотором интервале. Отметим следующие свойства производных высшего порядка.

1) **Постоянный множитель** можно выносить за знак n -й производной. Действительно,

$$(Cu)' = Cu',$$

$$(Cu)'' = (Cu')' = Cu'',$$

и вообще,

$$(Cu)^{(n)} = Cu^{(n)}.$$

¹⁾ Нетрудно заметить, что функции a^x , $\sin x$, $\cos x$ имеют производные любого порядка на всей числовой оси, $-\infty < x < \infty$. Функция x^α (α — любое действительное число) имеет производные любого порядка в области определения самой функции x^α , кроме быть может (в зависимости от α), точки $x=0$. Например, для функции $f(x) = x^{\frac{5}{3}}$, определенной на всей числовой оси, производная $f'(x) = \frac{5}{3} x^{\frac{2}{3}}$ определена

(существует) при всех x , а $f''(x) = \frac{10}{9} x^{-\frac{1}{3}}$ определена (существует) при всех x , отличных от нуля, т. е. в точке $x=0$ вторая производная этой функции не существует.

2) Производная n -го порядка алгебраической суммы функций равна алгебраической сумме производных n -го порядка этих функций.

Действительно,

$$(u \pm v)' = u' \pm v', \\ (u \pm v)'' = (u' \pm v')' = u'' \pm v''$$

и вообще

$$(u \pm v)^{(n)} = u^{(n)} \pm v^{(n)}.$$

3) Производная произведения.

$$(uv)' = u'v + uv', \\ (uv)'' = (u'v + uv')' = u''v + u'v' + u'v' + uv'' = u''v + 2u'v' + uv'', \\ (uv)''' = (u''v + 2u'v' + uv'')' = u'''v + u''v' + 2u''v' + 2u'v'' + \\ + u'v'' + uv''' = u'''v + 3u''v' + 3u'v'' + uv''.$$

Нетрудно заметить, что в каждом из равенств

$$(uv)' = u'v + uv', \\ (uv)'' = u''v + 2u'v' + uv'', \\ (uv)''' = u'''v + 3u''v' + 3u'v'' + uv''$$

коэффициенты такие же, как и у двучлена, возведенного соответственно в первую, вторую, третью степени, а сумма порядков производных в каждом слагаемом одна и та же и равна порядку производной uv .

В общем случае

$$(uv)^{(n)} = u^{(n)}v + nu^{(n-1)}v' + \frac{n(n-1)}{2!}u^{(n-2)}v'' + \dots + \\ + \frac{n(n-1)\dots[n-(k-1)]}{k!}u^{(n-k)}v^{(k)} + \dots + uv^{(n)}.$$

Эта формула n -й производной произведения двух функций называется *формулой Лейбница*.

Пример. $y = \frac{x^2}{1+x}$. Найти $y^{(V)}$.

Заметим, что при решении можно использовать следующий порядок записи:

$$\begin{array}{l|l|l} 1 & u = x^2 & v^{(V)} = -1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 (1+x)^{-6} \\ 5 & u' = 2x & v^{(IV)} = 1 \cdot 2 \cdot 3 \cdot 4 (1+x)^{-5} \\ 10 & u'' = 2 & v''' = -1 \cdot 2 \cdot 3 (1+x)^{-4} \\ 10 & u''' = 0 & v'' = 1 \cdot 2 (1+x)^{-3} \\ 5 & u^{(IV)} = 0 & v' = -(1+x)^{-2} \\ 1 & u^{(V)} = 0 & v = (1+x)^{-1} \end{array}$$

Тогда $y^{(V)} = -5!x^2(1+x)^{-6} + 10 \cdot 4!x(1+x)^{-5} - 20 \cdot 3!(1+x)^{-4}$.

И окончательно

$$y^{(V)} = -\frac{5!x^2}{(1+x)^6} + \frac{10 \cdot 4!x}{(1+x)^5} - \frac{20 \cdot 3!}{(1+x)^4}.$$

Дифференциалы высших порядков. Пусть функция $y = f(x)$ имеет производные n -го порядка на некотором интервале.

Рассмотрим дифференциал этой функции: $dy = f'(x) dx$, где $dx = \Delta x$ — приращение независимой переменной. Если фиксировать некоторое значение Δx , то dy будет функцией только независимой переменной x , определенной на рассматриваемом интервале.

Дифференциал от dy называется *дифференциалом второго порядка* (или вторым дифференциалом) функции $y = f(x)$, он обозначается d^2y («де два игрек»).

¹⁾ Нетрудно видеть, что обе функции y и $y^{(V)}$ определены на промежутках $(-\infty, -1)$ и $(-1, +\infty)$.

Итак, по определению $d^2y = d(dy)$, откуда, учитывая выражение дифференциала функции, запишем:

$$d^2y = (dy)' dx = [f'(x) dx]' dx = f''(x) dx^2.$$

Дифференциалом третьего порядка функции $y = f(x)$ называется дифференциал от дифференциала второго порядка этой функции, он обозначается d^3y . По определению $d^3y = (d^2y)' dx = (f''(x) dx^2)' dx = f'''(x) dx^3$,

$$d^3y = f'''(x) dx^3,$$

и вообще дифференциал n -го порядка функции $y = f(x)$ определяется как дифференциал от дифференциала порядка $(n-1)$ этой функции: $d^n y = d(d^{n-1}y)$.

При этом $d^n y = f^{(n)}(x) dx^n$.

Из полученного соотношения следует, что n -я производная функции равна отношению ее дифференциала n -го порядка к n -й степени дифференциала независимой переменной, и, следовательно, введенный ранее символ n -й производной $\frac{d^n y}{dx^n}$ становится реальной дробью — отношением дифференциала $d^n y$ к n -й степени dx^1 .

Заметим еще, что дифференциалы высших порядков ($n \geq 2$) не обладают уже свойством инвариантности, как это имело место для дифференциала первого порядка.

Действительно, пусть $y = f(x)$, $x = \varphi(t)$, t — независимая переменная, т. е. y — сложная функция аргумента t . Кроме того, $\varphi(t)$ — дважды дифференцируемая функция на некотором интервале изменения t , а $f(x)$ — дважды дифференцируемая функция на соответствующем интервале изменения x .

Тогда по определению

$$d^2y = [f(x)]'_t \cdot dt^2 = [f'(x) \varphi'(t)]_t dt^2 = [f''(x) (\varphi'(t))^2 + f'(x) \varphi''(t)] dt^2,$$

откуда, $d^2y = f''(x) dx^2 + f'(x) d^2x$. Если же x — независимая переменная, то $d^2x = (x)'' dx^2 = 0$ и $d^2y = f''(x) dx^2$.

Раздел 3

ТЕОРЕМЫ О СРЕДНЕМ

§ 13. Теорема Ролля. Теоремы Лагранжа и Коши

Теорема Ролля²⁾. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, дифференцируема на интервале (a, b) и значения функции на концах отрезка равны, $f(a) = f(b)$, то на интервале (a, b) существует точка ξ , $a < \xi < b$, в которой производная функции $f(x)$ равна нулю, $f'(\xi) = 0$.

Геометрический смысл теоремы Ролля: если условия теоремы выполняются, то на интервале (a, b) существует такая точка ξ , что в соответствующей ей точке кривой $y = f(x)$ касательная параллельна оси OX (рис. 186).

Заметим, что точек, в которых производная функции $f(x)$ равна нулю, может быть и несколько, теорема утверждает существование по крайней мере одной такой точки.

¹⁾ Учитывая связь между производной n -го порядка и дифференциалом n -го порядка функции $f(x)$, часто вместо слов «функция, имеющая производную n -го порядка» говорят, что функция n раз дифференцируема. Если n -ая производная функции $f(x)$ непрерывна на данном промежутке, говорят, что функция $f(x)$ на этом промежутке n раз непрерывно дифференцируема.

²⁾ Ролль (1652—1719) — французский математик.

Доказательство. По свойству функций, непрерывных на отрезке, функция $f(x)$ на отрезке $[a, b]$ принимает наибольшее и наименьшее значения, которые обозначим соответственно через M и m . Имеются две возможности: или $M = m$, или $M \neq m$. Рассмотрим их.

Пусть $M = m$. Тогда функция $f(x)$ на отрезке $[a, b]$ сохраняет постоянное значение и, следовательно, в любой точке интервала (a, b) ее производная равна нулю. В этом случае за ξ можно взять любую точку интервала (a, b) .

Пусть теперь $M \neq m$. Так как значения функции на концах отрезка равны, то хотя бы одно из значений M или m функция принимает внутри отрезка $[a, b]$. Пусть это будет, например, M . Обозначим через ξ , $a < \xi < b$, точку, в которой $f(\xi) = M$.

Докажем, что $f'(\xi) = 0$.

Действительно, так как M — наибольшее значение функции, то для любого Δx ¹⁾

$$\Delta f(\xi) = f(\xi + \Delta x) - f(\xi) \leq 0.$$

При этом

$$\frac{\Delta f(\xi)}{\Delta x} \begin{cases} \leq 0, & \text{если } \Delta x > 0, \\ \geq 0, & \text{если } \Delta x < 0. \end{cases} \quad (1)$$

Но, так как по условию производная в точке ξ существует, то, следовательно, существует $\lim_{\Delta x \rightarrow 0} \frac{\Delta f(\xi)}{\Delta x}$.

Учитывая, что

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x > 0}} \frac{\Delta f(\xi)}{\Delta x} \leq 0, \quad \text{а} \quad \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x < 0}} \frac{\Delta f(\xi)}{\Delta x} \geq 0,$$

заключаем, что $\lim_{\Delta x \rightarrow 0} \frac{\Delta f(\xi)}{\Delta x} = 0$, т. е. $f'(\xi) = 0$, что и требовалось доказать.

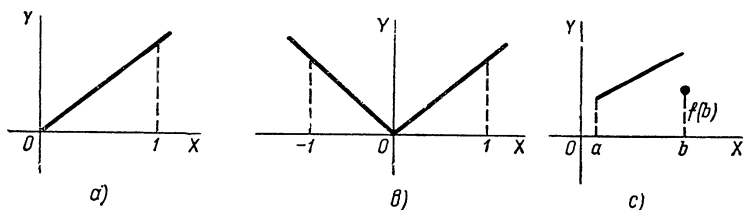


Рис. 187.

Заметим, что каждое из условий теоремы является необходимым. Это подтверждается рассмотрением функций, графики которых изображены на рис. 187. [а) $f(0) \neq f(1)$; б) $f(x)$ не дифференцируема в точке $x=0$; в) $f(x)$ не является непрерывной на отрезке $[a, b]$].

Нулем функции называется значение аргумента, при котором функция равна нулю.

Например, если $f(x) = \ln x$, то $x=1$ — нуль этой функции; если $f(x) = (x+2)^2$, то $x=-2$ — нуль функции.

Следствия теоремы Ролля

1) Если функция $f(x)$ на отрезке $[a, b]$ удовлетворяет условиям теоремы Ролля, причем $f(a) = f(b) = 0$, то существует по крайней мере одна точка ξ , $a < \xi < b$, такая, что $f'(\xi) = 0$. Другими словами, между двумя нулями функции находится по крайней мере один нуль ее производной.

¹⁾ Рассматриваем достаточно малые Δx , такие, чтобы точка $\xi + \Delta x$ принадлежала интервалу (a, b) .

2) Если на рассматриваемом выше интервале (a, b) функция $f(x)$ имеет производные $(n-1)$ -го порядка и n раз обращается в нуль, то существует по крайней мере одна точка интервала, в которой производная $(n-1)$ -го порядка равна нулю. В самом деле, по следствию из теоремы Ролля функция $f'(x)$ обращается в нуль по крайней мере в $(n-1)$ -й точке интервала (a, b) , $f''(x)$ обращается в нуль по крайней мере в $(n-2)$ -х точках этого интервала и т. д., $f^{(n-1)}(x)$ имеет по крайней мере один нуль на рассматриваемом интервале.

Теорема Лагранжа. Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема на интервале (a, b) , то на этом интервале найдется по крайней мере одна точка ξ , $a < \xi < b$, такая, что

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

Или, что то же, если на некотором отрезке выполняются условия теоремы, то отношение приращения функции к приращению аргумента на этом отрезке равно значению производной в некоторой промежуточной точке.

Заметим, что теорема Ролля является частным случаем теоремы Лагранжа при условии, что $f(a) = f(b)$.

Геометрический смысл теоремы Лагранжа.

Отношение $\frac{f(b) - f(a)}{b - a}$ равно угловому коэффициенту секущей AB (рис. 188).

Если функция $f(x)$ удовлетворяет условиям теоремы, то на интервале (a, b) существует точка ξ , такая, что в соответствующей точке кривой $y = f(x)$ касательная параллельна секущей, соединяющей точки $A(a, f(a))$, $B(b, f(b))$ (рис. 188). Таких точек может быть и несколько, но, по крайней мере, одна всегда существует.

Доказательство. Рассмотрим вспомогательную функцию

$$F(x) = f(x) - y_{\text{сек. } AB}.$$

Уравнение секущей AB [по точке $A(a, f(a))$ и угловому коэффициенту] имеет вид:

$$y - f(a) = \frac{f(b) - f(a)}{b - a}(x - a),$$

при этом

$$F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Функция $F(x)$ удовлетворяет условиям теоремы Ролля. Действительно, она непрерывна на отрезке $[a, b]$ и дифференцируема на интервале (a, b) . Нетрудно проверить, подставляя $x = a$ и $x = b$, что $F(a) = 0$ и $F(b) = 0$, т. е. $F(a) = F(b)$.

По теореме Ролля существует по крайней мере одна такая точка ξ , $a < \xi < b$, что $F'(\xi) = 0$. Так как

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

то, следовательно, $F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0$. Откуда получаем:

$f'(\xi) = \frac{f(b) - f(a)}{b - a}$, что и требовалось доказать.

Равенство

$$f(b) - f(a) = f'(\xi)(b - a),$$

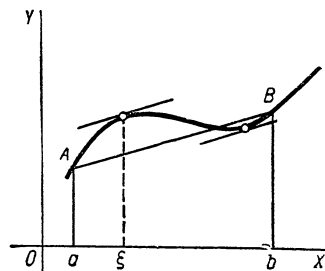


Рис. 188.

$a < \xi < b$, называется *формулой Лагранжа*, или *формулой конечных приращений*.

Так как точка ξ лежит между a и b (рис. 189), то можно записать $\xi = a + \theta(b - a)$, где $0 < \theta < 1$.

Здесь $\theta(b - a)$ является частью длины отрезка $[a, b]$. Учитывая это, формулу Лагранжа можно записать в виде:

$$f(b) - f(a) = f'(a + \theta(b - a))(b - a), \quad (2)$$

где $0 < \theta < 1$.

Заметим, что полученная запись формулы Лагранжа справедлива и в случае $b < a$. В самом деле, по доказанному имеем:

$$f(a) - f(b) = f'(\xi)(a - b), \quad (1')$$

где $\xi = a + \theta(b - a)$, $0 < \theta < 1$ (если учесть, что $b - a < 0$) (рис. 190).

Умножая равенство (1') на -1 и подставляя вместо ξ ее выражение, получим:

$$f(b) - f(a) = f'(a + \theta(b - a))(b - a), \quad 0 < \theta < 1.$$

Таким образом, формула Лагранжа (2) справедлива как при $a < b$, так и при $a > b$.

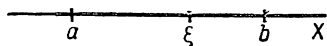


Рис. 189.

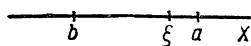


Рис. 190.

В частности, положив $a = x$, $b = x + \Delta x$, получим равенство $f(x + \Delta x) - f(x) = f'(x + \theta \Delta x) \Delta x$ или, что то же,

$$\Delta y = f'(x + \theta \Delta x) \Delta x. \quad (3)$$

И, следовательно, $\frac{\Delta y}{\Delta x} = f'(x + \theta \Delta x)$, где $0 < \theta < 1$.

Таким образом, отношение $\frac{\Delta y}{\Delta x}$ равно значению производной в некоторой смещенной точке.

Задача. Доказать справедливость неравенства $\frac{x}{x+1} \leq \ln(1+x) \leq x$.

Решение. Прежде всего отметим, что функция $\ln(1+x)$ существует при $x > -1$. Применяя формулу Лагранжа к $f(x) = \ln(1+x)$ на отрезке $[0, x]$, получим:

$$\ln(1+x) - \ln 1 = \frac{x}{1+\theta x} \quad \text{или} \quad \ln(1+x) = \frac{x}{1+\theta x} \quad (0 < \theta < 1).$$

Но $\frac{x}{1+x} \leq \frac{x}{1+\theta x} \leq x$ (равенство имеет место только при $x=0$). Окончательно получим: $\frac{x}{x+1} \leq \ln(1+x) \leq x$.

Теорема Коши¹⁾. Если функции $f(x)$ и $\varphi(x)$ непрерывны на отрезке $[a, b]$, дифференцируемы на интервале (a, b) и $\varphi'(x) \neq 0$ на интервале (a, b) , то существует по крайней мере одна точка ξ , $a < \xi < b$, такая, что

$$\frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)}, \quad (4)$$

т. е. отношение приращений функций на данном отрезке равно отношению их производных в точке ξ .

Прежде чем непосредственно перейти к доказательству теоремы Коши, сначала докажем, что $\varphi(b) - \varphi(a) \neq 0$.

¹⁾ Коши (1789—1857) — французский математик.

Действительно, если допустить, что $\varphi(b) - \varphi(a) = 0$ или, что то же, $\varphi(a) = \varphi(b)$, то по теореме Ролля для функции $\varphi(x)$ найдется точка η , $a < \eta < b$, в которой $\varphi'(\eta) = 0$. А это противоречит условию, так как $\varphi'(x) \neq 0$ на интервале (a, b) .

Доказательство теоремы. Рассмотрим вспомогательную функцию

$$F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} [\varphi(x) - \varphi(a)].$$

Нетрудно видеть, что эта функция на отрезке $[a, b]$ удовлетворяет условиям теоремы Ролля.

Действительно, функция $F(x)$ непрерывна на отрезке $[a, b]$, дифференцируемая на интервале (a, b) , кроме того, подстановка $x = a$ и $x = b$ дает $F(a) = 0$, $F(b) = 0$, т. е. $F(a) = F(b)$. По теореме Ролля для функции $F(x)$ существует точка ξ , $a < \xi < b$, такая, что $F'(\xi) = 0$.

Так как $F'(x) = f'(x) - \frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} \varphi'(x)$, то

$$F'(\xi) = 0 = f'(\xi) - \frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} \varphi'(\xi).$$

Отсюда, учитывая, что $\varphi'(\xi) \neq 0$, получаем:

$$\frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)},$$

что и требовалось доказать.

Доказанное равенство (4) называется *формулой Коши*. Теорему Лагранжа можно рассматривать как частный случай теоремы Коши (случай, когда $\varphi(x) = x$).

Может показаться, что формулу Коши проще получить из формулы Лагранжа следующим образом: записать равенство (1) сначала для функций $f(x)$, а затем для функции $\varphi(x)$ и разделить первое равенство на второе. Такое рассуждение ошибочно, так как точки ξ в этих равенствах, вообще говоря, различны.

§ 14. Раскрытие неопределенностей

Пусть функции $f(x)$ и $\varphi(x)$ в точке a обращаются в нуль, $f(a) = \varphi(a) = 0$. Тогда отношение $\frac{f(x)}{\varphi(x)}$ теряет смысл при $x = a$, но предел отношения $\frac{f(x)}{\varphi(x)}$ при $x \rightarrow a$ может существовать. Как уже говорилось выше (гл. III, § 8), отыскание этого предела называется раскрытием неопределенности вида $\frac{0}{0}$.

Следующая теорема, называемая *правилом Лопиталья*¹⁾, практически облегчает задачу нахождения этого предела.

Правило Лопиталья. Если функции $f(x)$ и $\varphi(x)$ дифференцируемы вблизи a , непрерывны в точке a , $\varphi'(x)$ отлична от нуля вблизи a и $f(a) = \varphi(a) = 0$, то предел отношения функций при $x \rightarrow a$ равен пределу отношения их производных, если последний (конечный или бесконечный) существует:

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{\varphi'(x)}. \quad (1)$$

Требования непрерывности $f(x)$ и $\varphi(x)$ в самой точке a и равенства $f(a) = \varphi(a) = 0$ можно заменить условиями: $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} \varphi(x) = 0$. Тогда, доопределив $f(x)$ и $\varphi(x)$ в точке a , исходя из условий $f(a) = \lim_{x \rightarrow a} f(x) = 0$, $\varphi(a) = \lim_{x \rightarrow a} \varphi(x) = 0$, получим функции $f(x)$ и $\varphi(x)$, непрерывные в точке a и равные в этой точке нулю.

¹⁾ Лопиталь (1661—1704) — французский математик.

Доказательство. Применяя формулу Коши, получим:

$$\frac{f(x) - f(a)}{\varphi(x) - \varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)},$$

где ξ между a и x . Или, учитывая, что $f(a) = \varphi(a) = 0$,

$$\frac{f(x)}{\varphi(x)} = \frac{f'(\xi)}{\varphi'(\xi)}. \quad (2)$$

Пусть при $x \rightarrow a$ отношение $\frac{f'(x)}{\varphi'(x)}$ стремится к некоторому пределу. Так как ξ лежит между a и x , то при $x \rightarrow a$ получим $\xi \rightarrow a$, а следовательно, и отношение $\frac{f'(\xi)}{\varphi'(\xi)}$ стремится к тому же пределу. Учитывая равенство (2), получим: если при $x \rightarrow a$ существует предел отношения $\frac{f'(x)}{\varphi'(x)}$, то

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{\varphi'(x)},$$

а это и требовалось доказать¹⁾.

Правило Лопиталья можно применять и для отыскания одностороннего предела. В этом случае достаточно потребовать, чтобы функции $f(x)$ и $\varphi(x)$ удовлетворяли условиям теоремы, соответственно, в правой (левой) окрестности точки a .

Если же окажется, что $f'(x)$ и $\varphi'(x)$ бесконечно малые при $x \rightarrow a$, то можно ставить вопрос о повторном применении правила Лопиталья. Иногда приходится правило Лопиталья применять несколько раз.

Пример 1. $\lim_{x \rightarrow 0} \frac{\sin 7x}{\lg 2x} = \lim_{x \rightarrow 0} \frac{7 \cos 7x}{2 \sec^2 2x} = \frac{7}{2}.$

Пример 2. $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = \lim_{x \rightarrow 0} \frac{e^x}{1} = 1.$

Пример 3. $\lim_{x \rightarrow 0} \frac{e^x - e^{-x} - 2x}{x - \sin x} = \lim_{x \rightarrow 0} \frac{e^x + e^{-x} - 2}{1 - \cos x} = \lim_{x \rightarrow 0} \frac{e^x - e^{-x}}{\sin x} = \lim_{x \rightarrow 0} \frac{e^x + e^{-x}}{\cos x} = 2$

Докажем, что правило Лопиталья остается справедливым и при условии, что a — один из символов ∞ , $+\infty$, $-\infty$.

Предположим, что функции $f(x)$ и $\varphi(x)$ дифференцируемы вблизи a , $\varphi'(x)$ отлична от нуля вблизи a и $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} \varphi(x) = 0$. Если перечисленные

¹⁾ Если при $x \rightarrow a$ предел отношения $\frac{f'(x)}{\varphi'(x)}$ не существует, то правило Лопиталья неприменимо. При этом предел отношения функций может существовать. Например,

$$\lim_{x \rightarrow 0} \frac{x^2 \sin \frac{1}{x}}{\sin x} = \lim_{x \rightarrow 0} \frac{x \sin \frac{1}{x}}{\frac{\sin x}{x}} = 0,$$

т. е. предел отношения функций при $x \rightarrow 0$ существует. Предел отношения производных этих функций не существует. Действительно, отношение производных равно

$$\frac{2x \sin \frac{1}{x} - \cos \frac{1}{x}}{\cos x}. \text{ Первое слагаемое } \frac{2x \sin \frac{1}{x}}{\cos x} \text{ при } x \rightarrow 0 \text{ имеет предел, равный нулю.}$$

Второе слагаемое $\frac{\cos \frac{1}{x}}{\cos x}$ при $x \rightarrow 0$ не имеет предела, так как $\cos \frac{1}{x}$ не имеет предела. Убедиться в последнем предоставляем читателю.

выше условия выполняются, то предел отношения функций при $x \rightarrow a$ равен пределу отношения их производных, если последний (конечный или бесконечный) существует:

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{\varphi'(x)}.$$

Действительно, пусть, например, a — символ ∞ (рассуждения сохраняются и для символов $+\infty$, $-\infty$), положив $x = \frac{1}{y}$, получим, что $y \rightarrow 0$ при $x \rightarrow \infty$, и поэтому если существует предел отношения функций $\frac{f(x)}{\varphi(x)}$ при $x \rightarrow \infty$, то существует и предел отношения функций $\frac{f(\frac{1}{y})}{\varphi(\frac{1}{y})}$ при $y \rightarrow 0$ и они равны. То же можно сказать и об отношении их производных.

Функции $f(\frac{1}{y})$, $\varphi(\frac{1}{y})$ вблизи точки $y=0$ удовлетворяют условиям доказанной теоремы, поэтому на основании равенства (1) имеем:

$$\lim_{y \rightarrow 0} \frac{f(\frac{1}{y})}{\varphi(\frac{1}{y})} = \lim_{y \rightarrow 0} \frac{f'(\frac{1}{y}) \left(-\frac{1}{y^2}\right)}{\varphi'(\frac{1}{y}) \left(-\frac{1}{y^2}\right)} = \lim_{y \rightarrow 0} \frac{f'(\frac{1}{y})}{\varphi'(\frac{1}{y})},$$

откуда получим: $\lim_{x \rightarrow \infty} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{\varphi'(x)}$.

Неопределенность вида $\frac{\infty}{\infty}$. Если $\lim_{x \rightarrow a} f(x) = \infty$ (или $+\infty$, $-\infty$), $\lim_{x \rightarrow a} \varphi(x) = \infty$ (или $+\infty$, $-\infty$), где a — число или один из символов ∞ , $+\infty$, $-\infty$, то отношение $\frac{f(x)}{\varphi(x)}$ при $x \rightarrow a$ может иметь предел. Нахождение такого предела называется раскрытием неопределенности вида $\frac{\infty}{\infty}$.

Если функции $f(x)$ и $\varphi(x)$ вблизи a дифференцируемы, $\varphi'(x)$ вблизи a отлична от нуля и $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} \varphi(x) = \infty$, то предел отношения функций при $x \rightarrow a$ равен пределу отношения их производных, если последний (конечный или бесконечный) существует:

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{\varphi'(x)}.$$

Доказательство этой теоремы мы опускаем.

Пример 1. $\lim_{x \rightarrow +\infty} \frac{\ln x}{x} = \lim_{x \rightarrow +\infty} \frac{\frac{1}{x}}{1} = 0.$

Пример 2. $\lim_{x \rightarrow +\infty} \frac{x^a}{a^x} \quad (a > 0, a > 1).$

$$\begin{aligned} \lim_{x \rightarrow +\infty} \frac{x^a}{a^x} &= \lim_{x \rightarrow +\infty} \frac{ax^{a-1}}{a^x \ln a} = \lim_{x \rightarrow +\infty} \frac{a(a-1)x^{a-2}}{a^x (\ln a)^2} = \\ &= \dots = \lim_{x \rightarrow +\infty} \frac{a(a-1) \dots [a-(n-1)] x^{a-n}}{a^x (\ln a)^n}, \end{aligned}$$

где n — наименьшее натуральное число, такое, что $a-n \leq 0$ и, следовательно,

$$\lim_{x \rightarrow +\infty} \frac{x^a}{a^x} = 0.$$

Это равенство означает, что показательная функция с основанием, большим единицы, при $x \rightarrow +\infty$ растет быстрее любой положительной степени x .

Другие виды неопределенностей. Раскрытие неопределенностей вида $0 \cdot \infty$ и $\infty - \infty$ можно свести к рассмотренным выше случаям неопределенностей $\frac{0}{0}$ и $\frac{\infty}{\infty}$.

Пусть, например, $f(x) \rightarrow 0$, $\varphi(x) \rightarrow \infty$ при $x \rightarrow a$ (a — число или один из символов, ∞ , $+\infty$, $-\infty$), тогда

$$f(x) \varphi(x) = \begin{cases} \frac{f(x)}{\frac{1}{\varphi(x)}} — \text{неопределенность вида } \frac{0}{0}, \\ \frac{\varphi(x)}{\frac{1}{f(x)}} — \text{неопределенность вида } \frac{\infty}{\infty}. \end{cases}$$

Пример. $\lim_{\substack{x \rightarrow 0 \\ x > 0}} x \ln x = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\ln x}{\frac{1}{x}} = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = 0.$

И, наконец, рассмотрим *неопределенности вида* 0^0 , 1^∞ , ∞^0 . Такие неопределенности встречаются при нахождении предела при $x \rightarrow a$ функции $y = [f(x)]^{\varphi(x)}$, $f(x) > 0$ вблизи a . Для нахождения предела при $x \rightarrow a$ такой функции достаточно найти предел при $x \rightarrow a$ функции $\ln y = \varphi(x) \ln f(x)$. Действительно, если $\lim_{x \rightarrow a} \ln y = A$, то $\lim_{x \rightarrow a} y = e^{A^1}$.

При отыскании предела $\ln y$ при $x \rightarrow a$ придется раскрывать неопределенность вида $0 \cdot \infty$, которая приводится к виду $\frac{0}{0}$ или $\frac{\infty}{\infty}$.

Пример. $\lim_{\substack{x \rightarrow 0 \\ x > 0}} x^x.$

Пусть $y = x^x$, тогда $\ln y = x \ln x$.

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} \ln y = \lim_{\substack{x \rightarrow 0 \\ x > 0}} x \ln x = 0$$

(как было получено в предшествующем примере). И, следовательно,

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} y = \lim_{\substack{x \rightarrow 0 \\ x > 0}} x^x = 1.$$

§ 15. Формула Тейлора²⁾

I. Формула Тейлора для многочлена. Пусть $P(x)$ — многочлен степени n :

$$P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n, \quad (a_n \neq 0)$$

и a — некоторое число. Многочлен $P(x)$ по степеням x всегда можно записать по степеням разности $x - a$. Действительно, положим $x - a = t$ (t — новая переменная), тогда $x = t + a$ и

$$P(x) = P(t + a) = a_0 + a_1(t + a) + a_2(t + a)^2 + \dots + a_n(t + a)^n.$$

Возводя слагаемые в правой части равенства в соответствующие степени и группируя члены с одинаковыми степенями t , запишем

$$P(t + a) = c_0 + c_1t + c_2t^2 + \dots + c_nt^n.$$

¹⁾ Если A — символ ∞ или $+\infty$, то $\lim_{x \rightarrow a} y = \infty$, если A — символ $-\infty$, то $\lim_{x \rightarrow a} y = 0$.

²⁾ Тейлор (1685—1731) — английский математик.

Откуда, возвратившись снова к переменной x , получим

$$P(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \dots + c_n(x-a)^n$$

— разложение многочлена $P(x)$ по степеням разности $x-a$.

Для нахождения коэффициентов c_k ($k=0, 1, \dots, n$) этого многочлена будем $P(x)$ последовательно n раз дифференцировать. При этом

$$\begin{aligned} P(x) &= c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \dots + c_n(x-a)^n, \\ P'(x) &= c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \dots + nc_n(x-a)^{n-1}, \\ P''(x) &= 1 \cdot 2 \cdot c_2 + 2 \cdot 3c_3(x-a) + \dots + (n-1)nc_n(x-a)^{n-2}, \\ P'''(x) &= 1 \cdot 2 \cdot 3 \cdot c_3 + \dots + (n-2)(n-1)nc_n(x-a)^{n-3}, \\ &\vdots \\ P^{(n)}(x) &= 1 \cdot 2 \cdot \dots \cdot nc_n. \end{aligned}$$

Полагая в этих равенствах $x=a$, получим:

$$P(a) = c_0, \quad P'(a) = c_1, \quad P''(a) = 2! c_2, \quad P'''(a) = 3! c_3, \quad \dots, \quad P^{(n)}(a) = n! c_n,$$

$$\text{откуда } c_0 = P(a), \quad c_1 = \frac{P'(a)}{1!}, \quad c_2 = \frac{P''(a)}{2!}, \quad c_3 = \frac{P'''(a)}{3!}, \quad \dots, \quad c_n = \frac{P^{(n)}(a)}{n!}.$$

Таким образом, коэффициенты многочлена выражаются через значения многочлена и его производных в точке a . Подставим их выражения в $P(x)$:

$$P(x) = P(a) + \frac{P'(a)}{1!}(x-a) + \frac{P''(a)}{2!}(x-a)^2 + \dots + \frac{P^{(n)}(a)}{n!}(x-a)^n.$$

Полученное равенство называется *формулой Тейлора для многочлена*. Формула Тейлора дает возможность многочлен $P(x)$ по степеням x записать по степеням разности $(x-a)$. Этой формулой удобно пользоваться при вычислении значений многочлена $P(x)$ для значений аргумента, близких к a . В этом случае, как легко видеть, начиная с некоторой степени $x-a$, слагаемыми, из-за их малости, можно пренебречь и тем самым упростить вычисление. При $a=0$ коэффициенты многочлена $P(x)$ по степеням x выражаются через значения $P(x)$ и его производных в точке 0. В этом случае формула Тейлора для многочлена имеет следующий вид:

$$P(x) = P(0) + \frac{P'(0)}{1!}x + \frac{P''(0)}{2!}x^2 + \dots + \frac{P^{(n)}(0)}{n!}x^n.$$

II. Формула Тейлора для произвольной функции. Пусть функция $f(x)$ в некотором интервале, содержащем точку a , имеет непрерывные производные до $(n+1)$ -го порядка включительно.

Ставится задача: представить функцию $f(x)$ на этом интервале в виде многочлена по степеням разности $x-a$ и найти ошибку такого приближенного представления.

$$\text{Рассмотрим многочлен } P(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Этот многочлен называется *многочленом Тейлора n -й степени для функции $f(x)$* . Обозначим разность $f(x) - P(x)$ через $R_n(x)$. Тогда

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x). \quad (1)$$

Формула (1) называется *формулой Тейлора для функции $f(x)$ по степеням разности $x-a$ или в окрестности точки a* , при этом функция $R_n(x)$ называется *остаточным членом формулы Тейлора*.

Существуют различные выражения остаточного члена $R_n(x)$, получим одно из них, наиболее часто встречающееся.

Будем искать остаточный член $R_n(x)$ в виде $R_n(x) = \frac{M}{(n+1)!} (x-a)^{n+1}$, где $M = M(x)$ — некоторая, пока неизвестная, функция от x . Обозначим значение $M(x)$, соответствующее значению $x=b$ (b — любая фиксированная точка из рассматриваемого интервала), через M_0 , тогда

$$R_n(b) = \frac{M_0}{(n+1)!} (b-a)^{n+1}.$$

И, следовательно,

$$f(b) = f(a) + \frac{f'(a)}{1!} (b-a) + \frac{f''(a)}{2!} (b-a)^2 + \dots + \frac{f^{(n)}(a)}{n!} (b-a)^n + \frac{M_0}{(n+1)!} (b-a)^{n+1}. \quad (2)$$

Для определения M_0 рассмотрим вспомогательную функцию

$$F(x) = f(x) - \left[f(a) + \frac{f'(a)}{1!} (x-a) + \frac{f''(a)}{2!} (x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!} (x-a)^n + \frac{M_0}{(n+1)!} (x-a)^{n+1} \right].$$

Непосредственная подстановка в выражение $F(x)$ $x=a$ и $x=b$ показывает, что $F(a)=0$ и $F(b)=0$ [второе равенство получается, если учесть равенство (1)].

Функция $F(x)$, как разность функции $f(x)$ и многочлена, имеет на рассматриваемом интервале производные до $(n+1)$ -го порядка включительно. Найдем эти производные.

$$\begin{aligned} F'(x) &= f'(x) - \left[f'(a) + \frac{f''(a)}{1!} (x-a) + \dots + \frac{f^{(n)}(a)}{(n-1)!} (x-a)^{n-1} + \frac{M_0}{n!} (x-a)^n \right], \\ F''(x) &= f''(x) - \left[f''(a) + \frac{f'''(a)}{1!} (x-a) + \dots + \frac{f^{(n)}(a)}{(n-2)!} (x-a)^{n-2} + \right. \\ &\quad \left. + \frac{M_0}{(n-1)!} (x-a)^{n-1} \right], \end{aligned} \quad (3)$$

.....

$$\begin{aligned} F^{(n)}(x) &= f^{(n)}(x) - \left[f^{(n)}(a) + \frac{M_0}{1!} (x-a) \right], \\ F^{(n+1)}(x) &= f^{(n+1)}(x) - M_0. \end{aligned}$$

Подставляя в выражения производных $x=a$, получим:

$$F'(a) = F''(a) = \dots = F^{(n)}(a) = 0.$$

Таким образом, функция $F(x)$ удовлетворяет следующим условиям:

$$F(a) = F(b) = 0 \text{ и } F'(a) = F''(a) = \dots = F^{(n)}(a) = 0.$$

Нетрудно показать, что существует точка ξ между a и b , такая, что $F^{(n+1)}(\xi) = 0$. Действительно, так как $F(a) = F(b)$, то по теореме Ролля существует между точками a и b точка ξ_1 , такая, что $F'(\xi_1) = 0$. Учитывая, что $F'(a) = 0$ и $F'(\xi_1) = 0$, на основании теоремы Ролля заключаем, что существует точка ξ_2 между точками a и ξ_1 , а следовательно, между точками a и b , такая, что $F''(\xi_2) = 0$, и так далее. Продолжая аналогичные рассуждения, заключаем, что между точками a и b существует точка ξ_n , в которой выполняется равенство $F^{(n)}(\xi_n) = 0$.

Так как $F^{(n)}(a) = 0$ и $F^{(n)}(\xi_n) = 0$, то существует такая точка ξ между a и ξ_n , а следовательно, между точками a и b , что $F^{(n+1)}(\xi) = 0$. Но $F^{(n+1)}(\xi) = f^{(n+1)}(\xi) - M_0$, и, следовательно, $M_0 = f^{(n+1)}(\xi)$. Подставляя полученное значение M_0 в равенство (2), получим:

$$\begin{aligned} f(b) &= f(a) + \frac{f'(a)}{1!} (b-a) + \frac{f''(a)}{2!} (b-a)^2 + \dots + \frac{f^{(n)}(a)}{n!} (b-a)^n + \\ &\quad + \frac{f^{(n+1)}(\xi)}{(n+1)!} (b-a)^{n+1}, \end{aligned} \quad (3)$$

где ξ — некоторая точка между a и b .

Учитывая произвольность выбора точки b , можно в равенстве (2) вместо точки b рассматривать произвольную точку x указанного интервала, и, записав вместо b число x , получим:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \quad (4)$$

где ξ — точка, лежащая между a и x .

Равенство (4) представляет собой *формулу Тейлора* [см. (1)] с *остаточным членом*:

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \quad (5)$$

где точка ξ лежит между a и x . Число ξ можно записать и так:

$$\xi = a + \theta(x-a), \quad 0 < \theta < 1.$$

Формула (5) записи остаточного члена формулы Тейлора называется *формой Лагранжа*.

Если в формуле Тейлора (4) положить $n=0$, то получим:

$$f(x) = f(a) + \frac{f'(\xi)}{1!}(x-a)$$

или $\frac{f(x)-f(a)}{x-a} = f'(\xi)$. Это равенство, как известно, является формулой Лагранжа.

Таким образом, формула Лагранжа является частным случаем формулы Тейлора.

Если же в формуле (4) положить $n=1$, то получим:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(\xi)}{2!}(x-a)^2$$

или, что то же,

$$f(x) - f(a) = \frac{f'(a)}{1!}(x-a) + \frac{f''(\xi)}{2!}(x-a)^2. \quad (6)$$

Обозначив $x-a = \Delta x$, получим $f(x) - f(a) = \Delta f(a)$, $f'(a)(x-a) = df(a)$ и, следовательно, на основании равенства (6): $\Delta f(a) \approx df(a)$; абсолютная погрешность такого приближения порядка $(\Delta x)^2$ (точнее, порядка не ниже $(\Delta x)^2$).

Учитывая выражения для дифференциалов различных порядков функции $f(x)$, можно записать формулу (4) в дифференциальной форме:

$$\Delta f(a) = \frac{df(a)}{1!} + \frac{d^2f(a)}{2!} + \dots + \frac{d^n f(a)}{n!} + \frac{d^{n+1} f(\xi)}{(n+1)!}. \quad (7)$$

И, наконец, если положить в формуле (4) $a=0$, то получим так называемую *формулу Маклорена с остаточным членом в форме Лагранжа*:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1}, \quad (4')$$

где $0 < \theta < 1$.

Пример 1. Написать разложение функции $f(x) = e^x$ по формуле Маклорена для любого n .

Решение. $f(x) = f'(x) = \dots = f^{(n+1)}(x) = e^x$; $f(0) = f'(0) = \dots = f^{(n)}(0) = 1$. Тогда

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \frac{e^\xi}{(n+1)!}x^{n+1}, \quad \text{где } \xi = \theta x, \quad 0 < \theta < 1.$$

Если заменить функцию e^x ее многочленом Тейлора степени n , то получим приближенное равенство:

$$e^x \approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}, \quad (8)$$

абсолютная погрешность которого равна

$$|R_n(x)| = \frac{e^\xi}{(n+1)!} |x|^{n+1}.$$

Если рассматривать функцию e^x для $-1 \leq x \leq 1$, то

$$|R_n(x)| \leq \frac{e^\xi}{(n+1)!} < \frac{3}{(n+1)!}.$$

(В этом случае $-1 < \xi < 1$). Полагая в формуле (8) $x=1$, получим приближенное значение числа e :

$$e \approx 1 + 1 + \frac{1}{2!} + \dots + \frac{1}{n!}.$$

При этом абсолютная погрешность меньше $\frac{3}{(n+1)!}$.

Если требуется вычислить значение числа e , например, с точностью до 0,01, то число n определяется из усиленного неравенства $\frac{3}{(n+1)!} < 0,001$ или $(n+1)! > 3000$. И, следовательно, если взять $n=6$, то требуемое неравенство удовлетворяется.

Пример 2. Написать разложение функции $f(x) = \sin x$ по формуле Маклорена.

Решение. Так как $f(x) = \sin x$, то $f'(x) = \cos x$, $f''(x) = -\sin x$, $f'''(x) = -\cos x$, $f^{IV}(x) = \sin x$, ... , $f^{(n)}(x) = \sin\left(x + n \frac{\pi}{2}\right)$ при любом n .

Полагая $x=0$, получаем $f(0)=0$; $f'(0)=1$; $f''(0)=0$; $f'''(0)=-1$; $f^{IV}(0)=0$, ...

Нетрудно заметить, что числа 0, 1, 0, -1 будут периодически повторяться, и что значения производных четных порядков при $x=0$ равны нулю.

Учитывая это, получим:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^{n+1} \frac{x^{2n-1}}{(2n-1)!} + R_{2n}(x),^1$$

где

$$R_{2n}(x) = \frac{f^{(2n+1)}(\xi)}{(2n+1)!} x^{2n+1} = \pm \frac{\cos \xi}{(2n+1)!} x^{2n+1}$$

и

$$|R_{2n}| \leq \frac{|x|^{2n+1}}{(2n+1)!}.$$

Пусть требуется определить значения x , при которых приближенное равенство $\sin x \approx x - \frac{x^3}{3!}$ справедливо с точностью до 0,0005. Так как $|R_n(x)| = \frac{|\cos \xi|}{5!} |x|^5$, то найдем x из усиленного неравенства

$$\frac{|x|^5}{5!} < 0,0005 \quad \text{или} \quad |x|^5 < 0,06,$$

откуда с помощью логарифмирования получаем $|x| < 0,57$.

Пример 3. Написать разложение функции $f(x) = \cos x$ по формуле Маклорена. Так как $f(x) = \cos x$, $f'(x) = -\sin x$, $f''(x) = -\cos x$, $f'''(x) = \sin x$, $f^{IV}(x) = \cos x$, ... , $f^{(n)}(x) = \cos\left(x + n \frac{\pi}{2}\right)$ при любом n , то $f(0)=1$; $f'(0)=0$; $f''(0)=-1$; $f'''(0)=0$; $f^{IV}(0)=1$; ...

¹⁾ В формулу (4') вместо n подставлено $2n$.

Нетрудно заметить, что числа 1, 0, —1, 0 будут периодически повторяться, при этом значения производных нечетных порядков при $x=0$ равны нулю. И, следовательно,

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + R_{2n+1}(x),^1$$

где

$$R_{2n+1}(x) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} x^{2n+2},$$

$$R_{2n+1}(x) = \pm \frac{\cos \xi}{(2n+2)!} x^{2n+2}.$$

Раздел 4

ИССЛЕДОВАНИЕ ФУНКЦИЙ И ПОСТРОЕНИЕ ГРАФИКОВ

§ 16. Возрастание и убывание функций

Функция $f(x)$ называется *возрастающей* на данном промежутке, если для любых двух точек x_1 и x_2 этого промежутка из неравенства $x_1 < x_2$ следует неравенство $f(x_1) < f(x_2)$ (рис. 191).

Из этого определения следует, что функция $f(x)$ возрастает на данном промежутке, если для любых двух точек x и $x + \Delta x$ этого промежутка знак приращения функции

$$\Delta y = f(x + \Delta x) - f(x)$$

совпадает со знаком Δx или, короче, если на данном промежутке приращение аргумента и соответствующее ему приращение функции — одного знака.

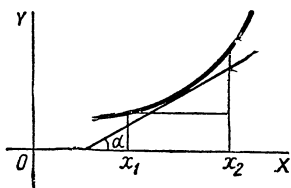


Рис. 191.

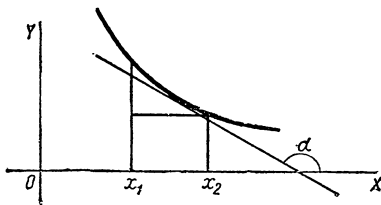


Рис. 192.

Функция $f(x)$ называется *убывающей* на данном промежутке, если для любых точек x_1 и x_2 этого промежутка из неравенства $x_1 < x_2$ следует неравенство $f(x_1) > f(x_2)$ (рис. 192). Другими словами, функция убывает на данном промежутке, если на этом промежутке приращение аргумента и соответствующее ему приращение функции разных знаков (при этом x и $x + \Delta x$ принадлежат рассматриваемому промежутку).

Необходимое условие возрастания и убывания дифференцируемой функции. Если дифференцируемая функция $f(x)$ возрастает на данном промежутке, то в любой точке x этого промежутка $f'(x) \geq 0$. Если дифференцируемая функция $f(x)$ убывает на данном промежутке, то в любой точке x этого промежутка $f'(x) \leq 0$.

Действительно, пусть, например, функция $f(x)$ возрастает на данном промежутке. Рассмотрим на нем произвольную точку x . Дадим

¹⁾ В формуле (4') вместо n взято $2n + 1$.

x приращение Δx , соответствующее ему приращение функции Δy одного знака с Δx , следовательно,

$$\frac{\Delta y}{\Delta x} > 0.$$

Переходя в этом неравенстве к пределу при $\Delta x \rightarrow 0$, получим:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(x) \geq 0.$$

Аналогично доказывается теорема и для убывающей функции $f(x)$.

Заметим, что если промежуток содержит какой-нибудь свой конец (или оба), то для x , совпадающего с этим концом, рассматривается односторонняя производная функции $f(x)$ (так как точки x и $x + \Delta x$ берутся из рассматриваемого промежутка).

Геометрически утверждение теоремы означает, что в каждой точке графика возрастающей функции касательная или образует острый угол с осью OX или параллельна оси OX , а в каждой точке графика убывающей функции касательная или образует тупой угол с осью OX или параллельна оси OX . Например, функция $y = x^3$ возрастает на всей числовой оси. Ее производная $y' = 3x^2$ положительна всюду, кроме точки $x = 0$, при $x = 0$ имеем $y' = 0$. Таким образом, касательная к кубической параболе в точке $x = 0$ параллельна оси OX , в остальных точках составляет острый угол с осью OX .

Достаточное условие возрастания и убывания функции. Пусть $f(x)$ — дифференцируемая функция на некотором промежутке.

Если в каждой точке x данного промежутка $f'(x) > 0$, то функция $f(x)$ возрастает на этом промежутке. Если в каждой точке x данного промежутка $f'(x) < 0$, то функция $f(x)$ убывает на этом промежутке.

Доказательство. Пусть, например, на данном промежутке $f'(x) > 0$. Возьмем на нем две любые точки x_1 и x_2 . Тогда по формуле Лагранжа

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1),$$

где ξ лежит между x_1 и x_2 . Так как $f'(\xi) > 0$, то знаки разностей $x_2 - x_1$ и $f(x_2) - f(x_1)$ совпадают, т. е. функция на данном промежутке возрастает.

Аналогично рассматривается случай, когда $f'(x) < 0$.

Интервалы, на которых функция возрастает (убывает), называются *интервалами монотонности* функции. По доказанному у дифференцируемой функции $f(x)$ на интервале возрастания $f'(x) \geq 0$, на интервале убывания $f'(x) \leq 0$. Если производная $f'(x)$ функции $f(x)$ непрерывна, то разделять интервалы монотонности могут лишь точки, в которых $f'(x) = 0$, так как перемена знака непрерывной функции возможна лишь при переходе ее через нуль. Точка, в которой $f'(x) = 0$, называется *точкой стационарности функции $f(x)$* , так как в достаточно малой окрестности этой точки функция «близка» к постоянной («почти» не изменяется). Заметим, что не каждая точка стационарности разделяет интервалы монотонности, например, функции $y = x^2$ и $y = x^3$ имеют точку стационарности $x = 0$. Эта точка для $y = x^2$ разделяет, а для $y = x^3$ не разделяет интервалы монотонности. Если не требовать непрерывности производной $f'(x)$ функции $f(x)$, то интервалы монотонности могут разделять не только точки стационарности $f(x)$. Например, для $f(x) = |x|$ точка $x = 0$ разделяет интервалы монотонности, в этой точке производная функции $f(x)$ не существует.

Ограничимся рассмотрением функций, дифференцируемых всюду, кроме, быть может, конечного числа точек, в их областях определения и имеющих не более конечного числа точек стационарности. Если $f(x)$ — функция рассмотренного класса, то можно доказать¹⁾, что ее интервалы

¹⁾ Доказательства этого положения не приводим.

монотонности разделяются или точками стационарности $f'(x)$, или точками, в которых производная функции $f'(x)$ не существует. Но не каждая такая точка будет разделять интервалы монотонности.

На рис. 193 отмечены точки $x_1, x_2, \dots, x_{10}$, в которых производная равна нулю или не существует. Причем точки $x_2, x_3, x_4, x_7, x_9, x_{10}$ разделяют, а точки x_1, x_5, x_6, x_8 не разделяют интервалов монотонности.

Учитывая сказанное, можно сформулировать **правило исследования функций указанного класса на возрастание и убывание**.

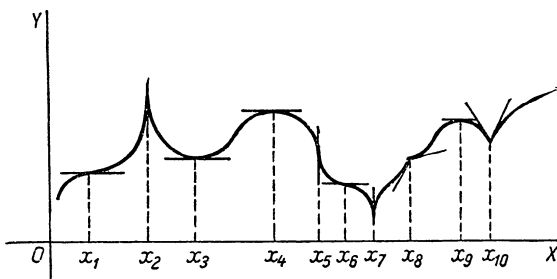


Рис. 193.

1) Находим точки из области определения функции, в которых производная функции равна нулю или не существует. Эти точки называют *критическими для функции $f(x)$ по первой производной*.

Критические точки разбивают область определения функции $f(x)$ на интервалы, на каждом из которых производная $f'(x)$ сохраняет знак.

2) Исследуем знак $f'(x)$ на каждом из этих интервалов.

Если на рассматриваемом интервале $f'(x) > 0$, то это интервал возрастания, если же $f'(x) < 0$, то это интервал убывания.

Пример 1. Найти интервалы монотонности функции $f(x) = x + \frac{1}{x}$.

Решение. Область определения функции состоит из интервалов $(-\infty, 0)$, $(0, +\infty)$, т. е. функция определена всюду, кроме точки $x = 0$; $f(x)$ — нечетная функция. Ее про-

изводная $f'(x) = 1 - \frac{1}{x^2}$ сущест-

вует в области определения функции $f(x)$, $f'(x) = 0$ при $x = \pm 1$.

Точки $-1, 0, 1$ разделяют область определения функции на интервалы, которые отмечены на рис. 194.

Учитывая нечетность функции $f(x)$, замечаем, что достаточно исследовать знак $f'(x)$ на интервалах $(0, 1)$ и $(1, +\infty)$.

$$f'(x) = \frac{(x-1)(x+1)}{x^2}.$$

Результат запишем в таблицу:

$0 < x < 1$	$f'(x) < 0$	$f(x)$ убывает
$1 < x < +\infty$	$f'(x) > 0$	$f(x)$ возрастает

Окончательно получаем: функция $f(x)$ возрастает на интервалах $(-\infty, -1)$, $(1, +\infty)$ и убывает на интервалах $(-1, 0)$, $(0, 1)$.

Пример 2. Исследовать на возрастание и убывание функцию

$$f(x) = x \ln x.$$

Решение. Функция $f(x)$ определена на интервале $(0, +\infty)$. Производная $f'(x) = \ln x + 1$ существует на этом интервале, $f'(x) = 0$ при $x = \frac{1}{e}$. Составим таблицу:

x	$f'(x)$	$f(x)$
$0 < x < \frac{1}{e}$	—	убывает
$\frac{1}{e} < x < +\infty$	+	возрастает

из которой видно, что $f(x)$ убывает на интервале $(0, \frac{1}{e})$ и возрастает на интервале $(\frac{1}{e}, +\infty)$.

§ 17. Экстремумы функции

Пусть функция $f(x)$ непрерывна на некотором интервале (a, b) , содержащем точку x_0 . Точка x_0 называется *точкой максимума* функции $f(x)$, если существует такая окрестность точки x_0 , что для всех $x (x \neq x_0)$ этой окрестности выполняется неравенство $f(x) < f(x_0)$. Точка x_0 называется *точкой минимума* функции $f(x)$, если существует такая окрестность точки x_0 , что для всех $x (x \neq x_0)$ этой окрестности выполняется неравенство $f(x) > f(x_0)$. Другими словами, x_0 — точка максимума (минимума) функции $f(x)$, если приращение функции в этой точке при всех достаточно малых приращениях аргумента отрицательно (положительно).

Точки максимума и минимума функции называют *точками экстремума*¹⁾ функции.

Значение функции $f(x)$ в точке максимума (минимума) называют *максимумом* (*минимумом*) функции, или *экстремумом* функции.

Функция $f(x)$ на данном промежутке может иметь и несколько экстремумов, причем некоторые из минимумов функции могут быть больше некоторых из ее максимумов.

На рис. 195 изображен график функции $y = f(x)$, у которой x_1 и x_3 — точки максимума, x_2 и x_4 — точки минимума. Минимум функции в точке x_4 больше ее максимума в точке x_1 .

Последнее не противоречит определению экстремума функции, так как в определении экстремума сравниваются значения функции в точке и некоторой ее окрестности.

Заметим, что точки максимума (минимума) функции $f(x)$ иногда определяют следующим образом: точка x_0 называется точкой максимума (минимума) функции $f(x)$, если существует такая окрестность точки x_0 , что для всех x этой окрестности $f(x) \leq f(x_0)$ ($f(x) \geq f(x_0)$).

В дальнейшем мы будем пользоваться ранее данным определением, т. е. рассматривать так называемый «строгий» экстремум. Заметим только, что все теоремы, которые будут доказаны для «строгого» экстремума, с небольшими изменениями легко можно перенести на случай «нестрогого» экстремума; предоставляем это сделать читателю.

¹⁾ «Экстремум» (extremum) означает по-латыни «крайний».

Необходимое условие экстремума дифференцируемой функции.

Теорема. Если дифференцируемая функция $f(x)$ имеет в точке x_0 экстремум, то $f'(x_0) = 0$.

Доказательство. Пусть, например, функция $y = f(x)$ имеет в точке x_0 максимум. Тогда $\Delta y = f(x_0 + \Delta x) - f(x_0) < 0$ при достаточно малом Δx , а следовательно,

$$\frac{\Delta y}{\Delta x} \begin{cases} < 0 & \text{при } \Delta x > 0, \\ > 0 & \text{при } \Delta x < 0. \end{cases}$$

Переходя к пределу при $\Delta x \rightarrow 0$, получим:

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x > 0}} \frac{\Delta y}{\Delta x} \leq 0; \quad \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x < 0}} \frac{\Delta y}{\Delta x} \geq 0.$$

По условию $f(x)$ — дифференцируемая функция в точке x_0 , следовательно, полученные пределы равны $f'(x_0)$. Таким образом, имеем $f'(x_0) \leq 0$ и $f'(x_0) \geq 0$, т. е. $f'(x_0) = 0$.

Заметим, что равенство нулю производной в точке x_0 является только необходимым условием экстремума дифференцируемой функции, но не достаточным, как это можно видеть на примере $y = x^3$. Здесь в точке $x = 0$ $y' = 0$, а функция $y = x^3$ в точке $x = 0$ не имеет экстремума.

Достаточные условия (первого порядка) существования экстремума.

Условимся говорить, что функция $f(x)$ меняет знак с плюса на минус при переходе через точку x_0 , если $f(x) > 0$ слева от точки x_0 , и $f(x) < 0$ справа от x_0 . При этом рассматриваются значения x из достаточно малой окрестности x_0 .

Аналогично определяется перемена знака функции с минуса на плюс при переходе через точку x_0 .

Теорема. Если функция $f(x)$ непрерывна в точке x_0 и некоторой ее окрестности, дифференцируема в рассматриваемой окрестности x_0 , кроме, быть может, самой точки x_0 , и ее производная $f'(x)$ при переходе через точку x_0 меняет знак с плюса на минус, то функция $f(x)$ имеет максимум в точке x_0 , если же производная при переходе через точку x_0 меняет знак с минуса на плюс, то в точке x_0 функция $f(x)$ имеет минимум.

Доказательство. Пусть, например, при переходе через точку x_0 производная $f'(x)$ меняет знак с плюса на минус. Тогда по формуле Лагранжа для любой точки x из рассматриваемой окрестности

$$f(x) - f(x_0) = f'(\xi)(x - x_0),$$

где ξ лежит между x и x_0 . Если $x < x_0$, то $f'(\xi) > 0$ и $f(x) - f(x_0) < 0$. Если же $x > x_0$, то $f'(\xi) < 0$, и мы получим, что $f(x) - f(x_0) < 0$.

Таким образом, в рассматриваемой окрестности точки x_0 выполняется неравенство $f(x) - f(x_0) < 0$, а это означает, что в точке x_0 функция $f(x)$ имеет максимум.

Аналогично рассматривается случай перемены знака $f'(x)$ с минуса на плюс.

Из приведенного доказательства следует, что если производная $f'(x)$ не меняет знака при переходе через точку x_0 , то в этой точке экстремума нет (так как не сохраняется знак разности $f(x) - f(x_0)$ в окрестности этой точки). Будем рассматривать функции класса, описанного на стр. 232. Нетрудно видеть, что для каждой такой функции точка макси-

муму отделяет интервал возрастания функции $f(x)$ ($f'(x) > 0$) от интервала ее убывания ($f'(x) < 0$), точка минимума отделяет интервал убывания от интервала возрастания. Таким образом, если функция исследована на возрастание и убывание, то точки экстремума выделяются автоматически.

В общем случае для *нахождения экстремума* функции рассматриваемого класса нужно: 1) Найти точки, принадлежащие области определения функции $f(x)$, в которых ее производная $f'(x)$ равна нулю или не существует, т. е. найти критические точки функции по первой производной. 2) Исследовать знак $f'(x)$ в некоторой окрестности каждой критической точки. При этом, если $f'(x)$ меняет знак при переходе через такую точку, то функция $f(x)$ в этой точке имеет экстремум. Если же знак $f'(x)$ не меняется при переходе через рассматриваемую точку, то функция $f(x)$ не имеет экстремума в этой точке.

Нетрудно видеть, что на рис. 193 критические точки x_5, x_6, x_8 не являются точками экстремума.

Пример 1. $f(x) = x \ln x$. Эта функция уже была исследована на возрастание и убывание (§ 16, пример 2). В точке $x = \frac{1}{e}$ рассматриваемая функция имеет минимум, равный $f\left(\frac{1}{e}\right) = -\frac{1}{e}$. Построим график $f(x) = x \ln x$. Сначала определим знаки

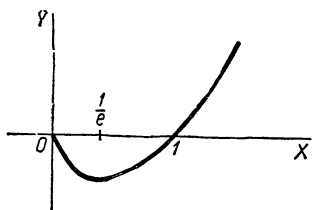


Рис. 196.

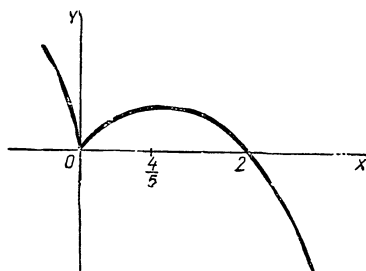


Рис. 197.

функции. Функция $f(x) = x \ln x$ равна нулю при $x = 1$. Область определения этой функции, $x > 0$, делится точкой $x = 1$ на интервалы $(0, 1)$, $(1, +\infty)$. При этом

$$\begin{aligned} f(x) &< 0, & \text{если} & \quad 0 < x < 1, \\ f(x) &> 0, & \text{если} & \quad 1 < x < +\infty. \end{aligned}$$

Найдем теперь пределы функции $f(x)$, если они существуют, при $x \rightarrow +\infty$ и $x \rightarrow +0$. Нетрудно видеть, что $\lim_{x \rightarrow +\infty} f(x) = +\infty$ и $\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} x \ln x =$

$$= \lim_{x \rightarrow 0} \frac{\ln x}{\frac{1}{x}} = \lim_{x \rightarrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = 0.$$

Суммируя все сказанное, построим график функции $f(x) = x \ln x$ (рис. 196).

Пример 2. $f(x) = x^{\frac{2}{3}}(2-x)$.

Функция $f(x)$ определена при всех значениях x :

$$f'(x) = \frac{2}{3}x^{-\frac{1}{3}}(2-x) - x^{\frac{2}{3}} = \frac{2(2-x) - 3x}{3\sqrt[3]{x}} = -\frac{5\left(x - \frac{4}{5}\right)}{3\sqrt[3]{x}},$$

$f'(x) = 0$ при $x = \frac{4}{5}$. Не существует производной функции $f(x)$ при $x = 0$.

Следовательно, критические точки функции $f(x)$: $x=0$ и $x=\frac{4}{5}$. Результат исследования запишем в таблицу:

x	$-\infty < x < 0$	$x=0$	$0 < x < \frac{4}{5}$	$x=\frac{4}{5}$	$\frac{4}{5} < x < +\infty$
y'	—	не существует	+	0	—
y	убывает	$f(0)=0$ минимум	возрастает	$f\left(\frac{4}{5}\right)=\frac{6}{5}\left(\frac{4}{5}\right)^{\frac{2}{3}}$ максимум	убывает

Заметим, что $f'(0)=+\infty$, $f'(-0)=-\infty$. Если учесть еще, что $f(x)=0$ при $x=0$ и $x=2$, и определить знак $f(x)$ в каждом из интервалов $(-\infty, 0)$, $(0, 2)$ и $(2, +\infty)$, а также учесть, что $f(x) \rightarrow \infty$ при $x \rightarrow \infty$, то график $f(x)$ изобразится так, как это показано на рис. 197.

Достаточные условия экстремума функции, выраженные через производные высших порядков. Предположим, что функция $f(x)$ в некоторой окрестности точки x_0 n раз непрерывно дифференцируема, т. е. имеет непрерывные производные до n -го порядка включительно.

Теорема. Пусть $f'(x_0)=f''(x_0)=\dots=f^{(n-1)}(x_0)=0$, а $f^{(n)}(x_0) \neq 0$. Тогда, если n — четное, то x_0 — точка экстремума функции [точка максимума при $f^{(n)}(x_0) < 0$ и точка минимума при $f^{(n)}(x_0) > 0$], если n — нечетное, то в точке x_0 экстремума нет.

Доказательство. Используя формулу Тейлора (§ 15), представим функцию $f(x)$ в окрестности точки x_0 в виде:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \dots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1} + \frac{f^{(n)}(\xi)}{n!}(x-x_0)^n,$$

где ξ лежит между x и x_0 .

Учитывая условия теоремы, получим

$$f(x) - f(x_0) = \frac{f^{(n)}(\xi)}{n!}(x-x_0)^n.$$

Так как функция $f^{(n)}(x)$ непрерывна в точке x_0 , то существует некоторая окрестность этой точки, в которой $f^{(n)}(x)$ сохраняет знак, т. е. знак $f^{(n)}(x)$ во всех точках этой окрестности совпадает со знаком $f^{(n)}(x_0)$. Если рассматривать значения x из достаточно малой окрестности точки x_0 , то знак $f^{(n)}(\xi)$ совпадает со знаком $f^{(n)}(x_0)$.

Пусть n — четное число, тогда множитель $(x-x_0)^n$ сохраняет знак в окрестности точки x_0 ($x \neq x_0$); при этом он положителен. Следовательно, разность $f(x) - f(x_0)$ сохраняет знак в окрестности x_0 ($x \neq x_0$), т. е. в точке x_0 функция $f(x)$ имеет экстремум. Кроме того, если $f^{(n)}(x_0) > 0$, то $f(x) - f(x_0) > 0$, т. е. x_0 — точка минимума функции $f(x)$, если же $f^{(n)}(x_0) < 0$, то $f(x) - f(x_0) < 0$ и в точке x_0 функция имеет максимум.

Пусть теперь n — нечетное число. Тогда множитель $(x-x_0)^n$ не сохраняет знака в окрестности точки x_0 ($x \neq x_0$) (при $x < x_0$ он отрицателен, а при $x > x_0$ он положителен), поэтому и разность $f(x) - f(x_0)$ не сохраняет знака в окрестности точки x_0 , а значит, функция в точке x_0 не имеет экстремума.

Частный случай теоремы ($n=2$). Если $f(x)$ — дважды непрерывно дифференцируемая функция в некоторой окрестности точки x_0 и если $f'(x_0)=0$, а $f''(x_0) \neq 0$, то x_0 является точкой экстремума функции. При

этом в точке x_0 функция имеет максимум, если $f''(x_0) < 0$ и минимум, если $f''(x_0) > 0$.

Рассмотрим п р и м е р. Найти экстремум функции $f(x) = (x-3)^2(x+2)^3$. Эта функция определена при всех значениях x

$$f'(x) = 2(x-3)(x+2)^3 + 3(x+2)^2(x-3)^2 = \\ = (x-3)(x+2)^2(2x+4+3x-9) = 5(x-3)(x+2)^2(x-1).$$

Производная $f'(x)$ существует при всех значениях x , поэтому экстремум функции $f(x)$ возможен только в нулях ее производной, т. е. в точках $x_1 = -2$, $x_2 = 1$, $x_3 = 3$.

Определим знаки $f''(x)$ в найденных точках стационарности. Так как $f''(x) = 5[(x+2)^2(x-1) + 2(x-3)(x+2)(x-1) + (x-3)(x+2)^2] = 5(x+2)[(x+2) \times (x-1) + 3(x-3)(x-1) + (x-3)(x+2)]$, то $f''(-2) = 0$ и, следовательно, вопрос об экстремуме в точке $x_1 = -2$ не может быть решен с помощью второй производной.

Для его решения нужно обратиться или к производным более высокого порядка в точке $x_1 = -2$ или провести исследование знака первой производной в некоторой окрестности точки -2 .

Так как $f'(x)$ не меняет знака при переходе через точку $x_1 = -2$, то в точке $x_1 = -2$ экстремума нет.

В точке $x_2 = 1$ $y''(1) < 0$, следовательно, в этой точке функция имеет максимум; при $x_3 = 3$ $y''(3) > 0$, в точке $x_3 = 3$ функция имеет минимум.

§ 18. О наибольшем и наименьшем значениях функции

На практике часто встречаются задачи, в которых требуется найти наибольшее или наименьшее значения функции на некотором промежутке. Будем решать эту задачу по отношению к функциям класса, описанного на стр. 232.

Известно, что если функция непрерывна на отрезке $[a, b]$, то она принимает на этом отрезке наибольшее (наименьшее) значение. В этом случае наибольшее (наименьшее) значение функции на этом отрезке равно или одному из значений, принимаемых функцией $f(x)$ в какой-нибудь критической точке (т. е. в точке, в которой производная $f'(x)$ равна нулю или не существует), принадлежащей интервалу (a, b) , или

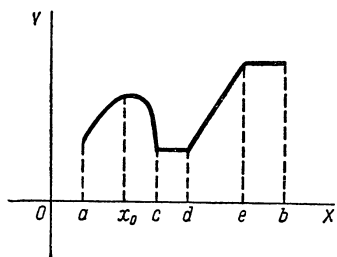


Рис. 198.

одному из значений, принимаемых этой функцией на концах отрезка. Заметим, что критическая точка функции $f(x)$, принадлежащая интервалу (a, b) , может не быть точкой экстремума этой функции даже в том случае, когда функция $f(x)$ принимает в ней свое наибольшее (наименьшее) на отрезке $[a, b]$ значение. На рис. 198 точки $x, c \leq x \leq d$, $e \leq x \leq b$, удовлетворяют указанному условию. При этом в точках $x, c \leq x \leq d$, функция $f(x)$ принимает наименьшее, а в точках $x, e \leq x \leq b$, наибольшее значение на отрезке $[a, b]$. Если функция непрерывна на

интервале, то она может не иметь на нем наибольшего (наименьшего) значения. Но если дифференцируемая функция на интервале (a, b) имеет не более конечного числа точек стационарности и только один экстремум — максимум (минимум), то наибольшее (наименьшее) значение функции на рассматриваемом интервале будет совпадать с максимумом (минимумом) этой функции.

Прежде чем перейти к рассмотрению задач на экстремум, сделаем несколько замечаний, упрощающих технику отыскания экстремумов.

1) Функции $f(x)$ и $f(x) + C$, где C — постоянное слагаемое, имеют экстремумы, и притом одноименные, в одних и тех же точках.

2) Функции $f(x)$ и $Cf(x)$, где C — постоянный множитель, имеют экстремумы в одних и тех же точках, причем одноименные, если $C > 0$, и разноименные, если $C < 0$.

3) Если на данном промежутке $f(x) > 0$, то функции $f(x)$ и $\sqrt{f(x)}$ имеют одноименные экстремумы в одних и тех же точках.

Задача 1. Из круга радиуса R (рис. 199) вырезан сектор, а из оставшейся части склеен конус. При каком угле вырезаемого сектора объем конуса будет наибольшим?

Решение. Образующая конуса равна R . Обозначив высоту и радиус основания соответственно через H и r , получим выражение объема конуса $V = \frac{1}{3} \pi r^2 H$.

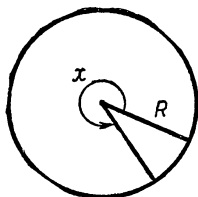


Рис. 199.

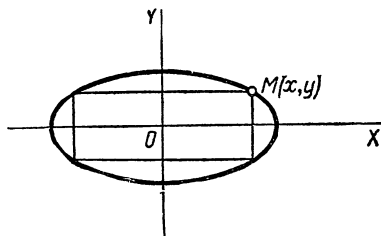
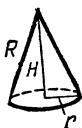


Рис. 200.

Пусть x — угол оставшегося (не вырезанного) сектора. Выразим r и H через x и R .

Длина окружности основания конуса равна $2\pi r = Rx$, откуда $r = \frac{Rx}{2\pi}$ и $H = \sqrt{R^2 - r^2} = \sqrt{R^2 - \frac{R^2 x^2}{4\pi^2}} = \frac{R}{2\pi} \sqrt{4\pi^2 - x^2}$.

Объем этого конуса определяется по формуле:

$$V = \frac{1}{3} \pi \frac{R^2 x^2}{4\pi^2} \cdot \frac{R}{2\pi} \sqrt{4\pi^2 - x^2} = \frac{R^3}{24\pi^2} x^2 \sqrt{4\pi^2 - x^2}.$$

Наибольшее значение объема (так как $V > 0$) достигается при тех же значениях x , что и наибольшее значение функции

$$f(x) = x^2 (4\pi^2 - x^2) = 4\pi^2 x^2 - x^4.$$

Заметив еще, что по условию задачи $0 < x < 2\pi$, найдем экстремумы функции $f(x)$ на этом интервале:

$$f'(x) = 8\pi^2 x - 4x^3 = 0 \quad \text{при} \quad x = 0; \quad x = \pm 2\pi \sqrt{\frac{2}{3}}.$$

На интервале $(0, 2\pi)$ лежит единственная точка стационарности $x = 2\pi \sqrt{\frac{2}{3}}$ функции $f(x)$.

Нетрудно убедиться в том, что функция $f(x)$, а следовательно, и объем V имеют в этой точке максимум. Это утверждение можно получить также и непосредственно, если учесть, что $f(0) = f(2\pi) = 0$, а $f(x) > 0$ на интервале $(0, 2\pi)$.

Так как точка экстремума одна и в ней объем имеет максимум, то наибольшее значение объема будет при $x = 2\pi \sqrt{\frac{2}{3}}$. И, следовательно, угол вырезаемого сектора должен быть равен $2\pi \left(1 - \sqrt{\frac{2}{3}}\right)$.

Задача 2. Из всех прямоугольников, вписанных в эллипс $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, найти тот, площадь которого наибольшая (предполагается, что стороны прямоугольника параллельны координатным осям, рис. 200).

Решение. Так как прямоугольник вписан в эллипс, то одна его вершина, обозначим ее $M(x, y)$, лежит в первой четверти. [Это утверждение следует хотя бы из того, что если вершина прямоугольника (x_0, y_0) лежит на эллипсе, то на эллипсе лежат и точки $(-x_0, y_0)$, $(x_0, -y_0)$, $(-x_0, -y_0)$ — остальные вершины прямоугольника.]

$$\text{Тогда площадь прямоугольника } S = 4xy = 4 \frac{b}{a} x \sqrt{a^2 - x^2} = \frac{4b}{a} \sqrt{a^2 x^2 - x^4}.$$

Очевидно, x изменяется на интервале $(0, a)$ и на этом интервале $S > 0$, поэтому для решения задачи достаточно найти наибольшее значение функции $f(x) = a^2 x^2 - x^4$ на рассматриваемом интервале.

Так как $f'(x) = 2ax - 4x^3$ и $f'(x) = 0$ при $x = 0$; $x = \pm \frac{a\sqrt{2}}{2}$, то на интервале $(0, a)$ находится единственная точка стационарности $x = \frac{a\sqrt{2}}{2}$, которая является точкой максимума функции $f(x)$ [хотя бы потому, что $f(a) = f(0) = 0$, $f(x) > 0$ на интервале $(0, a)$], а следовательно, и точкой максимума площади S .

В этой точке S имеет наибольшее значение.

Таким образом, прямоугольник с наибольшей площадью будет иметь размеры: $x = \frac{a\sqrt{2}}{2}$, $y = \frac{b\sqrt{2}}{2}$ и

$$S_{\text{наиб}} = 4 \frac{a\sqrt{2}}{2} \frac{b\sqrt{2}}{2} = 2ab.$$

Заметим, что наибольший по площади вписанный в эллипс прямоугольник имеет стороны, пропорциональные осям эллипса. При $a = b$ эллипс становится окружностью, следовательно, из всех прямоугольников, вписанных в окружность, наибольшую площадь имеет квадрат.

§ 19. Выпуклость и вогнутость. Точки перегиба

Любая прямая, не параллельная оси OY , разбивает плоскость на две области, из которых одна лежит выше, а другая ниже этой прямой (рис. 201).

Пусть в точке x_0 кривая $y = f(x)$ имеет касательную, не параллельную оси OY (т. е. $f(x)$ — дифференцируемая функция в точке x_0). Кривая называется *выпуклой* (*вогнутой*) в точке x_0 , если в некоторой окрестности этой точки кривая расположена ниже (выше) касательной, проведенной в точке x_0 .

На рис. 202 в точке x_1 кривая выпукла, в точке x_3 кривая вогнута, в точке x_2 кривая не является ни выпуклой, ни вогнутой.

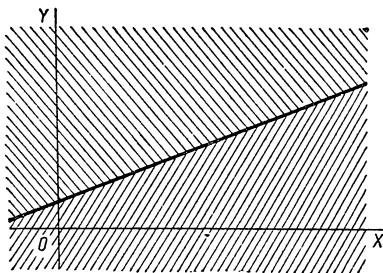


Рис. 201.

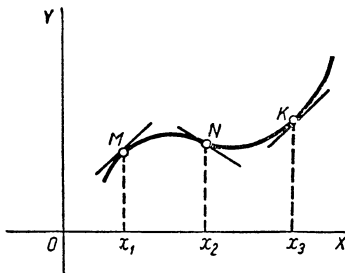


Рис. 202.

Кривая называется *выпуклой* (*вогнутой*) на интервале (a, b) , если она выпукла (вогнута) в каждой точке этого интервала.

Пусть кривая $y = f(x)$ в точке x_0 имеет касательную, в частности, эта касательная может быть и вертикальной. (В последнем случае функция $f(x)$ не дифференцируема в точке x_0 .)

Точка x_0 называется *точкой перегиба* кривой $y = f(x)$, если с одной стороны от x_0 (в достаточно малой окрестности точки x_0) кривая выпукла, с другой стороны от x_0 кривая вогнута или наоборот, если в точке x_0 кривая меняет выпуклость на вогнутость (или наоборот).

Из этого определения непосредственно следует, что касательная в точке перегиба кривой пересекает эту кривую. На рис. 202 x_2 — точка перегиба кривой, на рис. 203 точки x_1 , x_5 — точки перегиба, в точках x_2 , x_6 перегиба нет.

Заметим, что о точке перегиба кривой можно говорить лишь в случае, когда кривая имеет касательную в этой точке и в ее окрестности. (Например, точки x_3 , x_4 рис. 203 не являются точками перегиба).

Теорема. Если функция $f(x)$ в некоторой окрестности точки x_0 дважды непрерывно дифференцируема и $f''(x_0) \neq 0$, то необходимым и достаточным условием выпуклости (вогнутости) кривой $y=f(x)$ в точке x_0 является условие $f''(x_0) < 0$ ($f''(x_0) > 0$).

Доказательство. Запишем разложение функции $f(x)$ в окрестности x_0 по формуле Тейлора, взяв $n=1$:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(\xi)}{2!}(x-x_0)^2,$$

где ξ лежит между x и x_0 .

Как известно, касательная к графику функции $f(x)$ в точке x_0 имеет уравнение

$$y - f(x_0) = f'(x_0)(x - x_0).$$

Из этих двух равенств получаем:

$$y_{\text{кр}} - y_{\text{кас}} = \frac{f''(\xi)}{2!}(x - x_0)^2.$$

Знак разности $y_{\text{кр}} - y_{\text{кас}}$ совпадает со знаком $f''(\xi)$. Но по условию $f''(x)$ — непрерывная функция в точке x_0 и $f''(x_0) \neq 0$, следовательно, в достаточно малой окрестности точки x_0 функция $f''(x)$ сохраняет знак. Будем рассматривать эту окрестность точки x_0 , в ней знак $f''(\xi)$ совпадает со знаком $f''(x_0)$.

Учитывая, что $(x - x_0)^2 > 0$ при $x \neq x_0$, заключаем, что знак разности $y_{\text{кр}} - y_{\text{кас}}$ совпадает со знаком $f''(x_0)$. При этом, если $f''(x_0) < 0$, то $y_{\text{кр}} - y_{\text{кас}} < 0$ и, следовательно, кривая выпукла в точке x_0 , если $f''(x_0) > 0$, то $y_{\text{кр}} - y_{\text{кас}} > 0$ и, следовательно, кривая вогнута в точке x_0 . Наоборот, если кривая $y=f(x)$ выпукла в

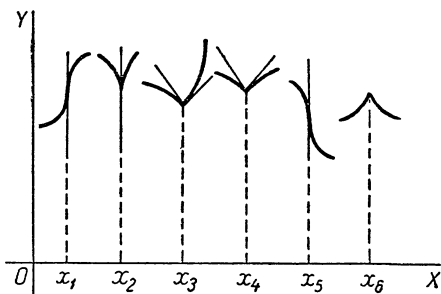


Рис. 203.

точке x_0 , то в некоторой окрестности этой точки $y_{\text{кр}} - y_{\text{кас}} < 0$ и, следовательно, $f''(x_0) < 0$. Если кривая вогнута в точке x_0 , то в некоторой ее окрестности $y_{\text{кр}} - y_{\text{кас}} > 0$ и, следовательно, $f''(x_0) > 0$.

Из доказанной теоремы непосредственно получаем необходимое условие точек перегиба кривой $y=f(x)$. Если $f(x)$ — дважды непрерывно дифференцируемая функция на некотором промежутке, то в точках перегиба кривой $y=f(x)$, лежащих на этом промежутке, вторая производная функции $f(x)$ равна нулю. Заметим, что не каждая такая точка является точкой перегиба кривой. Для функций x^3 и x^4 в точке $x=0$ производные $y''=6x$ и $y''=12x^2$ равны нулю. При этом для параболы $y=x^3$ точка $x=0$ является точкой перегиба, а парабола $y=x^4$ в точке $x=0$ перегиба не имеет.

Достаточное условие точки перегиба.

Теорема. Если в некоторой окрестности точки x_0 вторая производная $f''(x)$ непрерывна, всюду, кроме, может быть, самой точки x_0 (в ней $f''(x)$ может и не существовать), и при переходе через точку x_0 вторая производная меняет знак, то x_0 — точка перегиба кривой. (При этом предполагается, что кривая $y=f(x)$ имеет касательную в точке x_0 .)

Действительно, если знак второй производной при переходе через точку x_0 меняется с плюса на минус, то это означает, что вогнутость кривой сменилась на выпуклость при переходе через эту точку, т. е. x_0 — точка перегиба кривой. Аналогично точка x_0 является точкой перегиба, если знак $f''(x)$ изменился с минуса на плюс при переходе через точку x_0 . Заметим, что если $f''(x)$ не меняет знака при переходе через точку x_0 , то в рассматриваемой точке перегиба нет.

Ограничимся рассмотрением кривых $y=f(x)$, где функции $f(x)$ в области их определения дважды дифференцируемы, кроме, быть может, конечного числа точек, и $f''(x)$ имеют не более конечного числа нулей.

Тогда для кривой $y=f(x)$ рассматриваемого класса при условии, что в каждой точке она имеет касательную, можно доказать¹⁾, что ее интервалы выпуклости и вогнутости разделяются или нулями $f''(x)$, или точками, где $f''(x)$ не существует.

Для кривых $y=f(x)$ рассматриваемого класса можно сформулировать правила исследования на выпуклость, вогнутость и перегиб.

1) Находим точки, в которых $f''(x)=0$ или вторая производная функции $f(x)$ не существует. Такие точки называются *критическими точками функции $f(x)$ по второй производной*²⁾.

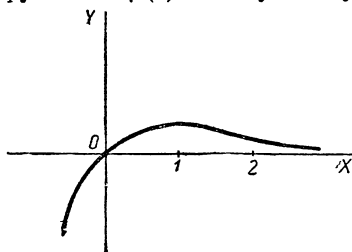


Рис. 204.

2) Эти точки разделят область определения функции $f(x)$ на интервалы, в каждом из которых $f''(x)$ сохраняет знак.

Если в рассматриваемом интервале $f''(x) > 0$, то это интервал вогнутости, если же $f''(x) < 0$, то это интервал выпуклости. Точки перегиба выделяются автоматически: они разделяют интервалы выпуклости и вогнутости. На рис. 203 графически иллюстрируются различные

возможные случаи. При этом только точки x_1 и x_2 являются точками перегиба.

Если же требуется найти только точки перегиба кривой $y=f(x)$, то можно исследовать знаки $f''(x)$ в окрестности каждой критической точки.

При этом окрестность рассматриваемой точки может быть как угодно большой, лишь бы в нее не попали другие точки, в которых возможен перегиб.

Пример 1. Найти интервалы выпуклости и вогнутости кривой

$$y = xe^{-x}.$$

Функция $f(x) = xe^{-x}$ определена при всех значениях x :

$$f'(x) = e^{-x} - xe^{-x} = (1-x)e^{-x},$$

$$f''(x) = -e^{-x}(1-x) - e^{-x} = -e^{-x}(2-x).$$

Вторая производная $f''(x)$ существует при всех значениях аргумента.

Найдем критические точки $f(x)$ по второй производной:

$$f''(x) = 0 \text{ при } x = 2.$$

Только в точке $x=2$ возможен перегиб. Результат исследования запишем в таблицу:

x	y''	y
$x < 2$	—	выпукла точка перегиба вогнута
$x = 2$	0	
$x > 2$	+	

Нетрудно проверить, что график этой функции имеет вид, изображенный на рис. 204.

Пример 2. Какие условия нужно наложить на коэффициенты a, b, c, d , чтобы кривая $y = ax^4 + bx^3 + cx^2 + dx + e$ имела точки перегиба?

Решение. $y'' = 12ax^2 + 6bx + 2c$.

Кривая имеет точку перегиба, если квадратное уравнение $6ax^2 + 3bx + c = 0$ имеет различные действительные корни, а это будет при условии, что $9b^2 - 24ac > 0$ или $3b^2 - 8ac > 0$.

¹⁾ Доказательство этого положения мы не приводим.

²⁾ При этом предполагается, что кривая в этих точках имеет касательную.

§ 20. Асимптоты плоских кривых

Понятие асимптоты уже встречалось при изучении гиперболы. Определим асимптоту кривой, заданной уравнением $y=f(x)$. Асимптотой кривой называется прямая, расстояние до которой от точки, лежащей на кривой, стремится к нулю при неограниченном удалении этой точки по кривой от начала координат или, что то же, когда расстояние точки от начала координат неограниченно растет. Не всякая кривая имеет асимптоту, например, эллипс асимптоты не имеет, парабола $y=x^2$ также не имеет асимптоты, так как при неограниченном удалении точки по параболе эта точка не может неограниченно приближаться к прямой (квадратичная функция растет быстрее любой линейной функции).

Пусть кривая $y=f(x)$ имеет асимптоту (рис. 205).

Предположим, что эта асимптота не вертикальная. Тогда ее уравнение имеет вид: $y=kx+b$. При этом $|y_{кр} - y_{кас}| = MN = \frac{MP}{\cos \varphi}$ ($\angle NMP = \varphi$). Так как угол наклона асимптоты постоянный, а $MP \rightarrow 0$ при неограниченном удалении точки M по кривой, т. е. при условии, что $x \rightarrow \infty$ (или $x \rightarrow +\infty$, или $x \rightarrow -\infty$), то $y_{кр} - y_{кас} = \alpha(x)$, где $\alpha(x) \rightarrow 0$ при $x \rightarrow \infty$ (или $x \rightarrow +\infty$, или $x \rightarrow -\infty$). Для упрощения в дальнейшем записанное в скобках будем опускать, имея в виду, что x может стремиться к бесконечности в одном или в двух направлениях. Таким образом, если $y=kx+b$ — асимптота кривой $y=f(x)$ при $x \rightarrow \infty$, то справедливо равенство

$$f(x) - (kx + b) = \alpha(x), \quad (1)$$

где $\alpha(x) \rightarrow 0$ при $x \rightarrow \infty$.

Обратно, если при каких-то постоянных k и b имеет место равенство (1), то $y=kx+b$ — асимптота кривой $y=f(x)$.

Пусть кривая $y=f(x)$ имеет асимптоту вида $y=kx+b$. Исходя из равенства (1), определим k и b . Для этого разделим равенство (1) на x и перейдем к пределу при $x \rightarrow \infty$.

Получим:

$$\frac{f(x)}{x} - \left(k + \frac{b}{x}\right) = \frac{\alpha(x)}{x},$$

$$\lim_{x \rightarrow \infty} \frac{f(x)}{x} - k = 0,$$

откуда

$$k = \lim_{x \rightarrow \infty} \frac{f(x)}{x}. \quad (2)$$

Теперь из равенства (1) определим b :

$$b = f(x) - kx - \alpha(x)$$

или, после перехода к пределу при $x \rightarrow \infty$:

$$b = \lim_{x \rightarrow \infty} [f(x) - kx]. \quad (3)$$

Наоборот, если существуют конечные пределы (2) и (3), то кривая $y=f(x)$ имеет асимптоту $y=kx+b$. Действительно, из равенства (3) следует равенство (1), на основании которого заключаем, что прямая $y=kx+b$ — асимптота кривой $y=f(x)$.

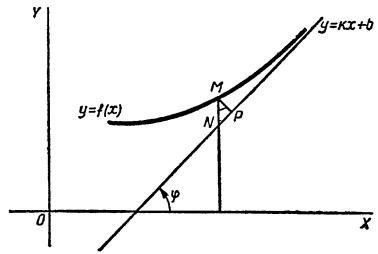


Рис. 205.

Итак, для того чтобы кривая $y=f(x)$ имела асимптоту $y=kx+b$, необходимо и достаточно, чтобы существовали конечные пределы (2) и (3).

Например, кривая $y=x+\sin x$ не имеет асимптоты при $x \rightarrow \infty$, так как не существует предела $\sin x$ при $x \rightarrow \infty$ ($k = \lim_{x \rightarrow \infty} \frac{x + \sin x}{x} = 1$).

Заметим, что если $k=0$ и b — конечно, то асимптота $y=b$ — горизонтальная (рис. 206).

Пусть теперь (рис. 207) кривая $y=f(x)$ имеет вертикальную асимптоту $x=a$ ($\varphi = \frac{\pi}{2}$). Тогда, как это следует из определения асимптоты, при $x \rightarrow a$ (или $x \rightarrow a$ с одной стороны) y неограниченно растет. [В этом случае расстояние точки $(x, f(x))$ кривой $y=f(x)$ до асимптоты $x=a$,

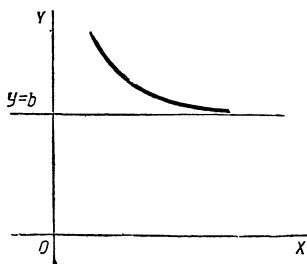


Рис. 206.

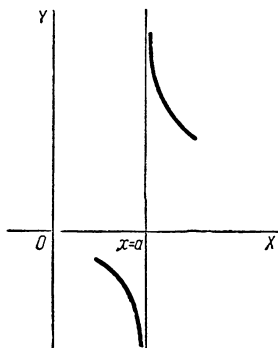


Рис. 207.

равное $|x-a|$, стремится к нулю]. Если существует число a , такое, что при $x \rightarrow a$ (или при $x \rightarrow a$ с одной стороны) $y \rightarrow \infty$, то $x=a$ — вертикальная асимптота.

Итак, если точка неограниченно удаляется по кривой $y=f(x)$, то эта кривая может иметь асимптоты в следующих трех случаях: $x \rightarrow \infty$, $y \rightarrow b$, тогда $y=b$ — горизонтальная асимптота; $x \rightarrow \infty$, $y \rightarrow \infty$, тогда возможна наклонная асимптота. И, наконец, $x \rightarrow a$, $y \rightarrow \infty$, в этом случае $x=a$ — вертикальная асимптота.

Рассмотрим примеры.

Пример 1. Найти асимптоты кривой $y = \frac{x^2 - 3x + 4}{x - 1}$.

Вертикальной асимптотой будет прямая $x=1$, так как при $x \rightarrow 1$ имеем $y \rightarrow \infty$.

Все остальные асимптоты (если они есть) имеют уравнения вида: $y=kx+b$. Найдем k и b :

$$k = \lim_{x \rightarrow \infty} \frac{f(x)}{x} = \lim_{x \rightarrow \infty} \frac{x^2 - 3x + 4}{x(x-1)} = 1$$

и

$$b = \lim_{x \rightarrow \infty} [f(x) - kx] = \lim_{x \rightarrow \infty} \left[\frac{x^2 - 3x + 4}{x-1} - x \right] = \lim_{x \rightarrow \infty} \frac{-2x + 4}{x-1} = -2,$$

$y=x-2$ — асимптота данной кривой при $x \rightarrow \infty$. Эта асимптота может быть получена также выделением целой части из рациональной дроби $\frac{x^2 - 3x + 4}{x-1}$.

Действительно, $y = \frac{x^2 - 3x + 4}{x-1} = x - 2 + \frac{2}{x-1}$, но $\frac{2}{x-1} \rightarrow 0$ при $x \rightarrow \infty$, следовательно, $y=x-2$ — асимптота.

Заметим, что если кривая имеет уравнение $y = \frac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ — многочлены, не имеющие общих корней, и многочлен $Q(x)$ имеет действительные корни $x_1, x_2, \dots, x_k$, то эта кривая имеет k вертикальных асимптот с уравнениями: $x=x_1, x=x_2, \dots, x=x_k$. У кривой $y = \frac{P(x)}{Q(x)}$ возможна также и горизонталь-

ная или наклонная асимптота при условии, что степень многочлена $P(x)$ не больше (случай горизонтальной асимптоты) или на единицу больше (случай наклонной асимптоты) степени многочлена $Q(x)$.

Пример 2. Найти асимптоты кривой $y = \frac{\ln x}{x}$. Функция $f(x) = \frac{\ln x}{x}$ определена при $x > 0$. Вертикальной асимптотой этой кривой будет $x = 0$, так как при $x \rightarrow 0$ ($x > 0$) $y \rightarrow -\infty$.

Будем искать другие асимптоты. Определим

$$k = \lim_{x \rightarrow +\infty} \frac{\ln x}{x^2} = 0, \quad b = \lim_{x \rightarrow +\infty} \frac{\ln x}{x} = 0,$$

таким образом, $y = 0$ — горизонтальная асимптота.

Заметим, что асимптота может пересекать кривую и даже в бесчисленном множестве точек. Например, кривая $y = \frac{\sin x}{x}$ (см. рис. 166) имеет асимптоту $y = 0$,

так как $\lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$.

Эта асимптота в точках $x = k\pi$ ($k = 0, \pm 1, \pm 2, \dots$) пересекает кривую.

§ 21. Построение графиков функций по критическим точкам и асимптотам

При построении графика функции можно использовать следующую схему:

- 1) Найти область определения функции.
- 2) Проверить функцию на четность и нечетность.
- 3) Исследовать знак функции $f(x)$ по ее критическим точкам, т. е. точкам, в которых $f(x) = 0$ или имеет разрыв.

- 4) Найти асимптоты графика функции.

- 5) Исследовать функцию на возрастание, убывание и экстремум.

- 6) Исследовать график функции на выпуклость, вогнутость и перегиб.

Не всегда нужно точно следовать этой схеме, в частности, пункт 2 в какой-то степени определяется пунктом 1, а пункт 6 используется не всегда. Иногда пунктов 1—4 достаточно для построения графика функции.

Кроме перечисленных пунктов, хорошо также посмотреть, как ведет себя функция при стремлении x к «граничным точкам» области ее определения, а также посмотреть, под каким углом касательная к графику функции (если она есть) идет в точках пересечения его с осями координат, в частности с осью OX , т. е. в нулях функции $f(x)$ и т. д.

Рассмотрим примеры на построение графиков функций.

Пример 1. $f(x) = \frac{x^3}{x^2 - 1}$.

Функция $f(x)$ определена при всех значениях x , кроме $x = \pm 1$; $f(x)$ — нечетная функция, так как $f(-x) = -f(x)$. Следовательно, для построения ее графика достаточно исследовать функцию только для $x \geq 0$. Будем рассматривать $x \geq 0$. Функция $f(x)$ равна нулю при $x = 0$; не существует при $x = 1$ и непрерывна в остальных точках, поэтому знаки $f(x)$ определяются следующим образом:

x	$0 < x < 1$	$1 < x < \infty$
$f(x)$	—	+

Следовательно, в интервале $(0, 1)$ график функции $f(x)$ расположен ниже, а в интервале $(1, +\infty)$ выше оси OX (заштриховываем те части плоскости, где нет кривой). Асимптоты: $x = -1$, $x = 1$ — вертикальные и $y = x$ — наклонная, так как $f(x) = x + \frac{x}{x^2 - 1}$ и $\frac{x}{x^2 - 1} \rightarrow 0$ при $x \rightarrow \infty$. Кроме того, график функции $f(x)$ при $x > 1$ выше асимптоты $y = x$ (так как $\frac{x}{x^2 - 1} > 0$ при $x > 1$), а на интервале $(0, 1)$ ниже асимптоты (так как $f(x) < 0$).

По этим данным уже можно представить характер графика функции $f(x)$. Нетрудно видеть, если x_0 ($x_0 > 0$) — точка максимума (минимума) функции $f(x) = \frac{x^3}{x^2-1}$, то точка, симметричная ей относительно начала координат, будет точкой минимума (максимума) этой функции. Найдем точки экстремума:

$$y' = \frac{3x^2(x^2-1) - 2x \cdot x^3}{(x^2-1)^2} = \frac{x^2(x^2-3)}{(x^2-1)^2}; \quad y' = 0 \quad \text{при} \quad x=0; \quad x=\pm\sqrt{3}.$$

Будем рассматривать только $x \geq 0$, тогда получим: $x=0$ — точка перегиба графика функции $f(x)$ (в силу симметрии графика относительно начала координат), в точке $x=\sqrt{3}$, как легко проверить, функция имеет максимум $f(\sqrt{3}) = \frac{3\sqrt{3}}{2}$. График этой функции изображен на рис. 208.

Пример 2. $f(x) = x^2 e^{-x}$. Эта функция определена на всей числовой оси, она неотрицательна и обращается в нуль только при $x=0$. Найдем асимптоты кривой

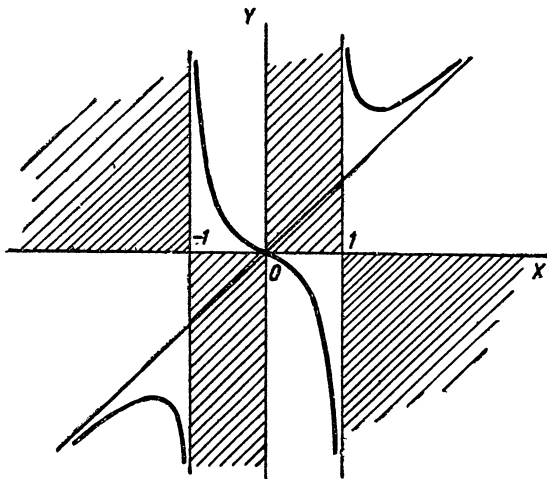


Рис. 208.

$y = x^2 e^{-x}$. Вертикальных асимптот эта кривая не имеет. Нетрудно видеть, что при $x \rightarrow -\infty$ кривая не имеет ни горизонтальной, ни наклонной асимптоты (так как $\lim_{x \rightarrow -\infty} f(x) = \infty$ и не существует конечного предела отношения $\frac{f(x)}{x}$ при $x \rightarrow -\infty$), при $x \rightarrow +\infty$ кривая имеет горизонтальную асимптоту $y=0$. (В этом случае $k = \lim_{x \rightarrow +\infty} \frac{x}{e^x} = 0$; $b = \lim_{x \rightarrow +\infty} \frac{x^2}{e^x} = 0$.)

Определим интервалы монотонности функции и исследуем функцию на экстремум.

$$y' = 2xe^{-x} - x^2 e^{-x} = x(2-x)e^{-x}.$$

Эта производная всюду существует и непрерывна; $y'=0$ при $x=0$ и $x=2$. Результат исследования запишем в таблицу.

x	$-\infty < x < 0$	$x=0$	$0 < x < 2$	$x=2$	$2 < x < +\infty$
y'	—	0	+	0	—
y	убывает	$f(0) = 0$ минимум	возрастает	$f(2) = 4e^{-2}$ максимум	убывает

И, наконец, исследуем кривую $y = x^2 e^{-x}$ на выпуклость и вогнутость и найдем ее точки перегиба:

$$y'' = (x^2 - 4x + 2)e^{-x},$$

y'' всюду существует и непрерывна. Так как $y''=0$ при $x=2 \pm \sqrt{2}$, то

$$y'' = [x - (2 - \sqrt{2})][x - (2 + \sqrt{2})]e^{-x}.$$

Составим таблицу:

x	$-\infty < x < 2 - \sqrt{2}$	$x = 2 - \sqrt{2}$	$2 - \sqrt{2} < x < 2 + \sqrt{2}$	$x = 2 + \sqrt{2}$	$2 + \sqrt{2} < x < +\infty$
y''	+	0	—	0	+
y	вогнута	$y = (2 - \sqrt{2})^2 e^{-(2 - \sqrt{2})}$ перегиб	выпукла	$y = (2 + \sqrt{2})^2 e^{-(2 + \sqrt{2})}$ перегиб	вогнута

На основании этой таблицы строим график функции $f(x) = x^2 e^{-x}$ (см. рис. 209).

Пример 3. $f(x) = \sqrt{8x^2 - x^4}$.

Область определения функции — совокупность значений x , удовлетворяющих условию $8x^2 - x^4 \geq 0$, откуда $x^2(8 - x^2) \geq 0$, $x^2 \leq 8$, $|x| \leq 2\sqrt{2}$. Следовательно, функция определена на отрезке $[-2\sqrt{2}, 2\sqrt{2}]$. Нетрудно видеть, что на этом

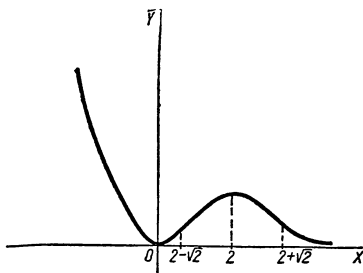


Рис. 209.

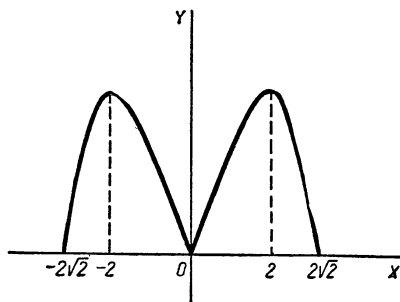


Рис. 210.

отрезке функция $f(x)$ неотрицательна, обращается в нуль лишь в точках $x=0$, $x=\pm 2\sqrt{2}$. В силу четности функции достаточно исследовать ее при $0 \leq x \leq 2\sqrt{2}$. Производная этой функции

$$y' = \frac{16x - 4x^3}{2\sqrt{8x^2 - x^4}} = -\frac{2x(x-2)(x+2)}{\sqrt{8x^2 - x^4}} = -\frac{2(x-2)(x+2)}{\sqrt{8-x^2}}.$$

Имеем $\lim_{x \rightarrow +0} y' = 2\sqrt{2}$. На интервале $(0, 2\sqrt{2})$ производная y' равна нулю только при $x=2$.левой производной при $x=2\sqrt{2}$ не существует, так как $\lim_{x \rightarrow 2\sqrt{2}-0} y' = -\infty$.

График функции $f(x)$ в точке $x=2\sqrt{2}$ имеет вертикальную касательную. В интервале $(0, 2\sqrt{2})$ имеется единственная точка $x=2$, в которой возможен экстремум.

Исследуем функцию на возрастание, убывание и экстремум.

x	$0 < x < 2$	$x = 2$	$2 < x < 2\sqrt{2}$
y'	+	0	—
y	возрастает	$f(2) = 4$ максимум	убывает

По этим данным можно построить график функции, он изображен на рис. 210.

§ 22. Параметрические уравнения кривых

Рассмотрим геометрическое место точек плоскости, координаты которых удовлетворяют уравнениям $x = \varphi(t)$, $y = \psi(t)$, где $\varphi(t)$ и $\psi(t)$ — непрерывные функции на некотором промежутке значений вспомогательного аргумента t , называемого *параметром*. Это геометрическое место точек называется *непрерывной кривой*, а записанные уравнения — *параметрическими уравнениями этой кривой*. Иногда от параметрического

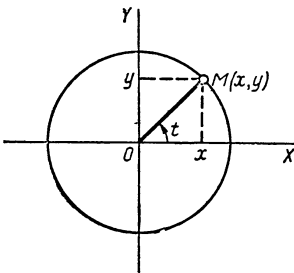


Рис. 211.

задания кривой можно перейти к заданию ее уравнением $y = f(x)$. Например, если для функции $x = \varphi(t)$ существует обратная функция $t = \theta(x)$, непрерывная в соответствующем интервале значений x , то $y = \psi[\theta(x)] = f(x)$ — непрерывная функция в рассматриваемом интервале значений x . Обратно, уравнение кривой $y = f(x)$ всегда можно, и не единственным способом, записать в параметрическом виде, в частности, положив $x = t$, получим $y = f(t)$.

Пример. Рассмотрим уравнение параболы $y = x^2$. Положив $x = t$, получим $y = t^2$ — совокупность этих двух уравнений, где $-\infty < t < +\infty$, можно рассматривать как параметрические уравнения параболы. Или, положив $x = t^3$, определим $y = t^6$ и, следовательно, $x = t^3$, $y = t^6$ — другие параметрические уравнения той же параболы, $-\infty < t < +\infty$.

Если кривая задана уравнением в полярных координатах $\rho = \rho(\theta)$, то, используя связь между прямоугольными и полярными координатами, получим $x = \rho(\theta) \cos \theta$, $y = \rho(\theta) \sin \theta$ — параметрические уравнения кривой, где роль параметра выполняет полярный угол θ . В механике обычно за параметр принимается время, в этом случае совокупность уравнений $x = x(t)$, $y = y(t)$, где t — время, называют уравнениями движения.

Примеры параметрических уравнений кривых.

1) Параметрические уравнения окружности. Пусть имеем окружность радиуса a с центром в начале координат (рис. 211). Возьмем за параметр t — центральный угол. Координаты любой точки этой окружности удовлетворяют уравнениям

$$\begin{cases} x = a \cos t, \\ y = a \sin t, \quad 0 \leq t \leq 2\pi, \end{cases}$$

которые, следовательно, и являются параметрическими уравнениями рассматриваемой окружности.

Исключая параметр t (возводя эти равенства в квадрат и складывая), получим $x^2 + y^2 = a^2$ — каноническое уравнение окружности.

2) Параметрические уравнения эллипса.

Рассмотрим две концентрические окружности с центром в начале координат радиусов a и b . Пусть t — центральный угол. Обозначим соответственно через N и K точки пересечения радиуса с окружностями (рис. 212) и через $M(x, y)$ точку пересечения прямой, проходящей через точку N параллельно оси OX , и прямой, проходящей через точку K параллельно оси OY .

Напишем параметрические уравнения геометрического места точек $M(x, y)$:

$$\begin{cases} x = a \cos t, \\ y = b \sin t, \quad 0 \leq t \leq 2\pi. \end{cases}$$

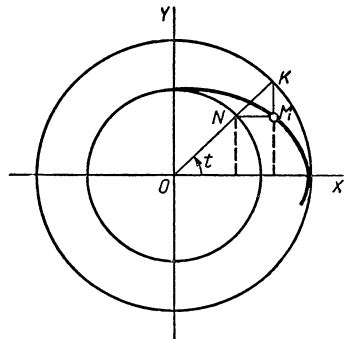


Рис. 212.

Рассматриваемое геометрическое место точек является эллипсом, так как, исключив параметр t из уравнений, получим: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ — каноническое уравнение эллипса.

3) **Параметрические уравнения циклоиды.** Циклоидой называется кривая, которую описывает некоторая точка, лежащая на окружности, когда окружность без скольжения катится по прямой.

Пусть окружность радиуса a катится по оси OX и в начальный момент выбранная точка M этой окружности находится в начале координат (рис. 213). Возьмем за параметр t угол поворота катящейся окружности:

$$t = \angle AC_1M.$$

Выразим координаты x и y точки M через параметр t :

$$x = \overline{OB} = \overline{OA} - \overline{BA} = at - a \sin t = a(t - \sin t),$$

аналогично получаем $y = a - a \cos t = a(1 - \cos t)$.

Итак,

$$\begin{cases} x = a(t - \sin t), \\ y = a(1 - \cos t), \end{cases} \quad -\infty < t < +\infty$$

— параметрические уравнения циклоиды.

4) **Параметрические уравнения пространственной кривой.** Геометрическое место точек пространства, координаты которых удовлетворяют уравнениям

$$x = f(t), \quad y = \varphi(t), \quad z = \psi(t) \dots (1),$$

где $f(t)$, $\varphi(t)$, $\psi(t)$ — непрерывные функции на некотором промежутке значе-

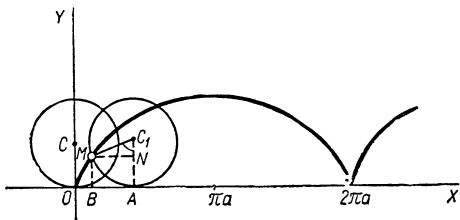


Рис. 213.

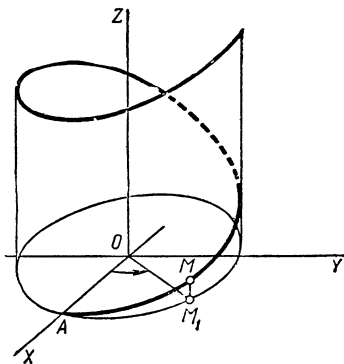


Рис. 214.

ний t , называется *непрерывной пространственной кривой*, а совокупность трех уравнений (1) — *параметрическими уравнениями этой кривой*.

В качестве примера рассмотрим параметрические уравнения винтовой линии.

Пусть окружность радиуса a с центром в начале координат лежит в плоскости XOY . Рассмотрим точку M этой окружности, находящуюся на положительной части оси OX .

Винтовой линией называется траектория точки M , равномерно движущейся по окружности, когда плоскость окружности равномерно перемещается в направлении, перпендикулярном плоскости XOY (рис. 214).

Так как проекция точки $M(x, y, z)$ на плоскость XOY равномерно движется по окружности, то, принимая за параметр t центральный угол $\angle AOM_1$, получим:

$$\begin{aligned} x &= a \cos t, \\ y &= a \sin t. \end{aligned}$$

Так как плоскость окружности равномерно перемещается в вертикальном направлении, то аппликата точки M пропорциональна t и третье параметрическое уравнение винтовой линии будет $z = ct$.

Итак, параметрические уравнения винтовой линии имеют вид:

$$x = a \cos t, \quad y = a \sin t, \quad z = ct \quad (-\infty < t < +\infty).$$

§ 23. Производные функций, заданных параметрически

Рассмотрим кривую с параметрическими уравнениями $x = \varphi(t)$, $y = \psi(t)$. Предположим, что уравнение этой кривой на некотором интервале изменения параметра t можно записать в виде: $y = f(x)$. Если $x = \varphi(t)$, $y = \psi(t)$ — дифференцируемые функции на рассматриваемом интервале значений t и $\varphi'(t) \neq 0$ в этом интервале, то можно показать¹⁾, что для $x = \varphi(t)$ существует обратная функция $t = \theta(x)$, дифференцируемая в соответствующем интервале изменения x , а значит, $y = \psi[\theta(x)] = f(x)$ — дифференцируемая функция на рассматриваемом интервале изменения x . Ставится задача — выразить производную y' функции $y = f(x)$ через производные x и y по параметру t .

Для решения этой задачи воспользуемся равенством $y' = \frac{dy}{dx}$, вытекающим из инвариантной формы записи дифференциала. Так как $dy = \psi'(t) dt$, $dx = \varphi'(t) dt$, то, следовательно, $y' = \frac{\psi'(t)}{\varphi'(t)}$ или $y' = \frac{\dot{y}}{\dot{x}}$, где через $\dot{y}$, $\dot{x}$ обозначены производные y и x по параметру t .

Формула $y' = \frac{\dot{y}}{\dot{x}}$ называется *формулой параметрического дифференцирования*.

Пользуясь этой формулой, можно при некоторых условиях на функции $\varphi(t)$, $\psi(t)$ получить выражение производной y по x любого порядка через производные x и y по t . Например,

$$y'' = (y')' = \left(\frac{\dot{y}}{\dot{x}} \right)' = \frac{\ddot{y}\dot{x} - \dot{y}\ddot{x}}{\dot{x}^3}.$$

Таким образом,

$$y'' = \frac{\ddot{y}\dot{x} - \dot{y}\ddot{x}}{\dot{x}^3}.$$

Пример. Найти y'' для функции: $\begin{cases} x = a(t - \sin t), \\ y = a(1 - \cos t), \end{cases}$

графиком которой является циклоида.

Решение. Подставляя в формулу второй производной выражения производных x и y по t , получим:

$$y'' = \frac{a(1 - \cos t) a \cos t - a \sin t \cdot a \sin t}{a^3 (1 - \cos t)^3} = -\frac{1}{a(1 - \cos t)^2} = -\frac{1}{4a} \operatorname{cosec}^4 \frac{t}{2}.$$

Пользуясь формулой параметрического дифференцирования, запишем уравнения касательной и нормали к кривой, заданной параметрическими уравнениями $x = \varphi(t)$, $y = \psi(t)$, в ее точке $M(x_0, y_0)$.

Как известно, уравнение невертикальной касательной к кривой, заданной уравнением $y = f(x)$, имеет вид:

$$y - y_0 = y'(x - x_0),$$

где $y' = f'(x_0)$ — значение производной в точке x_0 . Пусть точке $M(x_0, y_0)$ соответствует значение параметра $t = t_0$. Так как $y' = \frac{\dot{y}}{\dot{x}}$, то, следова-

¹⁾ Доказательство этого положения мы опускаем.

тельно, уравнение касательной при параметрическом задании кривой имеет вид:

$$y - y_0 = \frac{\dot{y}}{\dot{x}} (x - x_0),$$

т. е.

$$\frac{x - x_0}{\dot{x}} = \frac{y - y_0}{\dot{y}},$$

где $\dot{x}$, $\dot{y}$ — значения производных при $t = t_0$. И уравнение нормали в точке $M(x_0, y_0)$ к кривой, заданной в параметрическом виде, запишется в виде:

$$y - y_0 = -\frac{\dot{x}}{\dot{y}} (x - x_0),$$

т. е.

$$\dot{x}(x - x_0) + \dot{y}(y - y_0) = 0,$$

где $\dot{x}$ и $\dot{y}$ — значения производных при $t = t_0$.

Пример. Написать уравнение (в прямоугольной системе координат) касательной к логарифмической спирали $\rho = e^\theta$ в точке, для которой $\theta = 0$.

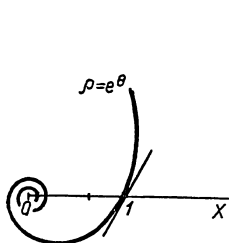


Рис. 215.

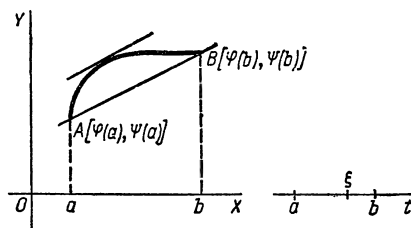


Рис. 216.

Решение. Параметрические уравнения логарифмической спирали записываются так: $x = e^\theta \cos \theta$, $y = e^\theta \sin \theta$. Точка касания M имеет прямоугольные координаты $(1, 0)$.

$$\begin{aligned} \dot{x} &= e^\theta (\cos \theta - \sin \theta), \\ \dot{y} &= e^\theta (\sin \theta + \cos \theta), \end{aligned}$$

при $\theta = 0$ $\dot{x} = \dot{y} = 1$.

Уравнение касательной будет $y = x - 1$ (см. рис. 215).

Если $f(x)$ и $\varphi(x)$ — функции, удовлетворяющие условиям теоремы Коши на отрезке $[a, b]$, то по формуле Коши

$$\frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)},$$

где ξ лежит между a и b .

Задача. Убедиться в том, что теорема Коши имеет такой же геометрический смысл, как и теорема Лагранжа (рис. 216), если рассмотреть кривую с параметрическими уравнениями $x = \varphi(t)$, $y = f(t)$, $a \leq t \leq b$.

§ 24. Уравнения касательной прямой и нормальной плоскости к пространственной кривой, заданной параметрически

Пусть точке $M(x_0, y_0, z_0)$ кривой с уравнениями $x = f(t)$, $y = \varphi(t)$, $z = \psi(t)$ соответствует значение параметра $t = t_0$. Дадим t приращение Δt , тогда $t_0 + \Delta t$ соответствует точка $N(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$. Уравнения секущей MN , как уравнения пространственной прямой, проходящей через две точки, можно записать так:

$$\frac{x - x_0}{\Delta x} = \frac{y - y_0}{\Delta y} = \frac{z - z_0}{\Delta z}.$$

Если разделить все знаменатели на Δt и перейти к пределу при $\Delta t \rightarrow 0$, то получим уравнения

$$\frac{x - x_0}{\dot{x}} = \frac{y - y_0}{\dot{y}} = \frac{z - z_0}{\dot{z}}$$

касательной к пространственной кривой, заданной параметрически. При этом $\dot{x}$, $\dot{y}$, $\dot{z}$ вычисляются при значении $t = t_0$, соответствующем точке касания $M(x_0, y_0, z_0)$. (Предполагается, что $\dot{x}$, $\dot{y}$, $\dot{z}$ существуют и не все равны нулю.)

Плоскость, перпендикулярная касательной к кривой и проходящая через точку касания, называется *нормальной плоскостью к кривой* в этой точке. Используя условие перпендикулярности прямой и плоскости, запишем уравнение нормальной плоскости к данной кривой в ее точке $M(x_0, y_0, z_0)$

$$\dot{x}(x - x_0) + \dot{y}(y - y_0) + \dot{z}(z - z_0) = 0.$$

Пример. Показать, что касательная к винтовой линии в любой ее точке образует постоянный угол γ с осью OZ .

Решение. $x = a \cos t$, $y = a \sin t$, $z = ct$ — уравнения винтовой линии. Тогда $\dot{x} = -a \sin t$, $\dot{y} = a \cos t$, $\dot{z} = c$ и уравнения касательной в любой точке винтовой линии запишутся в виде:

$$\frac{X - x}{-a \sin t} = \frac{Y - y}{a \cos t} = \frac{Z - z}{c},$$

где X, Y, Z — текущие координаты точки касательной. Известно, что

$$\cos \gamma = \frac{c}{\sqrt{(-a \sin t)^2 + (a \cos t)^2 + c^2}} = \frac{c}{\sqrt{a^2 + b^2}}.$$

Отсюда видно, что $\cos \gamma$ не зависит от точки касания. Это можно иллюстрировать следующим образом: если представить себе винтовую линию накрученной на прямой круглый цилиндр, то винтовая линия пересекает все образующие этого цилиндра под постоянным углом.

§ 25. Кривизна. Радиус кривизны

Прежде чем определить кривизну кривой в точке, как характеристику «степени искривленности» кривой, определим *положительное направление касательной* к кривой и *длину кривой*.

Рассмотрим кривую, в каждой точке которой имеется касательная и, следовательно, нормаль.

Предположим, что положение касательной непрерывно изменяется с изменением положения точки кривой. Это означает, что для любого числа $\epsilon > 0$ для каждой точки M кривой существует δ -окрестность, такая, что для всех точек кривой, попавших в эту окрестность, угол между нормальными к кривой в этих точках и в точке M меньше, чем ϵ . Такой кривой является, например, график непрерывно дифференцируемой функции $f(x)$ или кривая, заданная параметрическими уравнениями $x = \varphi(t)$, $y = \psi(t)$, где $\varphi(t)$ и $\psi(t)$ — непрерывно дифференцируемые функции параметра t и $\varphi'(t)$, $\psi'(t)$ одновременно не обращаются в нуль.

Выберем на кривой определенное «положительное» направление, т. е. установим, какую из любых двух точек M и N этой кривой считаем первой, а какую второй. Так как положение точки на рассматриваемой кривой однозначно определяется значением параметра в случае параметрического задания кривой (значением x , если кривая — график функции $f(x)$), то обычно за положительное направление на кривой принимают направление возрастания параметра (возрастание x). По условию в каждой точке рассматриваемой кривой имеется касательная. До сих пор мы не различали взаимно противоположные направления на самой касательной по той причине, что угловой коэффициент для этих двух направлений один и тот же. Иногда бывает нужно выбрать одно из двух возможных направлений касательной.

Пусть M — некоторая фиксированная точка рассматриваемой кривой. Возьмем с положительной стороны от точки M любую точку K этой кривой (т. е. точка M — первая, а K — вторая относительно выбранного направления) (рис. 217). Направление хорды MK , при стремлении точки K к точке M с положительной стороны, в пределе дает некоторое определенное направление касательной к кривой в точке M , которое называется *положительным*.

Пусть φ — угол между осью OX и положительным направлением касательной в какой-либо точке M кривой. При перемещении точки M по

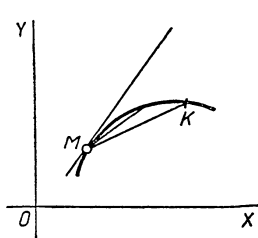


Рис. 217.

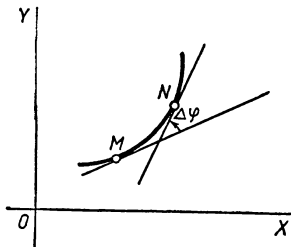


Рис. 218.

кривой φ будет изменяться. Приращение $\Delta\varphi$ угла φ при переходе на кривой от точки M к точке N называется *углом смежности* дуги MN рассматриваемой кривой (рис. 218).

Дифференциал дуги. Рассмотрим дугу кривой, заданной уравнением $y = f(x)$, где $f(x)$ — непрерывно дифференцируемая функция на некотором отрезке. И пусть $A(a, f(a))$, $B(b, f(b))$ — точки этой дуги. Будем считать направление от A к B положительным. Разобьем дугу AB на части точками $M_0, M_1, M_2, \dots, M_n$, занумерованными в направлении от A к B (здесь M_0 совпадает с A , M_n — с B). Соединив точки деления, получим ломаную $M_0M_1 \dots M_n$, вписанную в дугу $\overline{AB}$ (рис. 219). Дуга $\overline{AB}$ называется *спрямляемой*, если существует и конечен предел периметров вписанных в дугу $\overline{AB}$ ломаных при стремлении к нулю наибольшей из длин звеньев этих ломаных.

В противном случае дуга $\overline{AB}$ называется *неспрямляемой*.

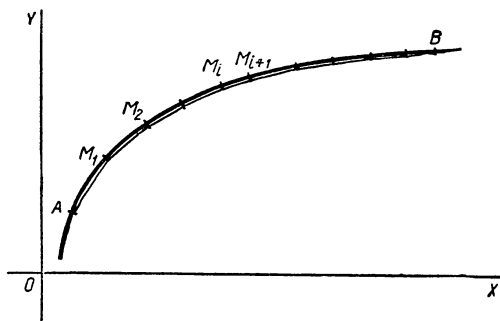


Рис. 219.

Если дуга $\overline{AB}$ спрямляема, то предел периметров вписанных в нее ломаных называется *длиной дуги $\overline{AB}$* . В § 12 гл. VI доказывается, что если $f(x)$ — непрерывно дифференцируемая функция на отрезке $[a, b]$, то дуга $\overline{AB}$ спрямляема, и дается формула для вычисления длины дуги. Там же доказывается, что предел отношения длины дуги к длине стягивающей ее хорды при стягивании дуги в точку равен единице, т. е. длины хорды и дуги — эквивалентные бесконечно малые. Длину дуги кривой будем обозначать через s . Тогда каждой точке M дуги можно поставить в соответствие число, которое будем обозначать также через s , равное длине дуги $\overline{AM}$, с определенным знаком: $s > 0$, если точка $M(x, f(x))$ лежит с положительной стороны от A (в данном случае $x > a$), $s < 0$,

если $M(x, f(x))$ лежит с отрицательной стороны от A (в данном случае $x < a$), и $s=0$, если M совпадает с A ¹⁾). Таким образом определяется возрастающая функция $s=s(x)$ (так как точка кривой однозначно определяется ее абсциссой). Найдем выражение для $ds(x)$. Дадим x приращение Δx , тогда $s(x)$ получит приращение $\Delta s = s(x + \Delta x) - s(x)$. Длины

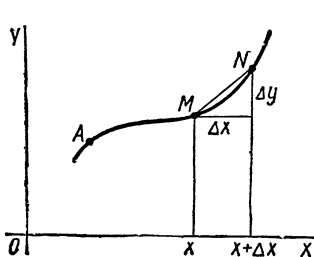


Рис. 220.

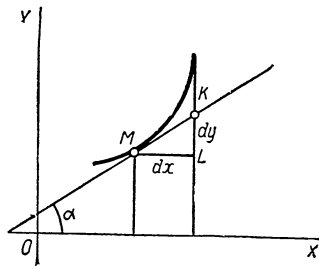


Рис. 221.

хорды MN и $|\Delta s|$ — эквивалентные бесконечно малые. Обозначая через $\overline{MN}$ длину отрезка MN , взятую со знаком плюс (минус), если N правее (левее) M , и учитывая, что Δs и направленная длина отрезка MN одного знака, получаем:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta s}{\overline{MN}} = 1$$

(на рис. 220 $\Delta x > 0$). Так как $MN = \sqrt{(\Delta x)^2 + (\Delta y)^2}$, кроме того, знаки $\overline{MN}$ и Δx совпадают, то

$$\overline{MN} = \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2} \cdot \Delta x.$$

Следовательно,

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta s}{\Delta x \cdot \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2}} = 1,$$

т. е.

$$\frac{ds}{dx} \cdot \frac{1}{\sqrt{1 + (y')^2}} = 1,$$

откуда получаем:

$$\frac{ds}{dx} = \sqrt{1 + (y')^2}, \quad ds = \sqrt{1 + (y')^2} dx,$$

или, наконец, $ds = \pm \sqrt{(dx)^2 + (dy)^2}$, при этом знак корня совпадает со знаком $dx = \Delta x$. Нетрудно видеть, что $|ds|$ геометрически представляет собой длину отрезка MK касательной к кривой в точке $M(x, f(x))$ (рис. 221), а запись $|ds| = \sqrt{(dx)^2 + (dy)^2}$ можно рассматривать как «теорему Пифагора» для дифференциалов. Из $\triangle MKL$ получаем:

$$\cos \alpha = \frac{dx}{ds} = \frac{1}{\sqrt{1 + (y')^2}}, \quad \sin \alpha = \frac{dy}{ds} = \frac{y'}{\sqrt{1 + (y')^2}},$$

при этом рассматривается положительное направление касательной.

Кривизна кривой. Радиус кривизны. Средней кривизной дуги $\widehat{MN}$ называется отношение угла смежности дуги к величине Δs этой дуги:

$$k_{\widehat{MN}} = \frac{\Delta \varphi}{\Delta s}.$$

¹⁾ В дальнейшем под s будем понимать направленную величину.

Средняя кривизна характеризует среднюю скорость изменения угла смежности дуги вдоль этой дуги.

Величина средней скорости на различных участках дуги, вообще говоря, будет различной.

Кривизной k кривой в точке M называется предел средней кривизны дуги $\widehat{MN}$ кривой при стремлении точки N по кривой к точке M .

$$k = \lim_{N \rightarrow M} \frac{\Delta\varphi}{\Delta s} = \lim_{\Delta s \rightarrow 0} \frac{\Delta\varphi}{\Delta s}$$

или, если φ — дифференцируемая функция переменной s ,

$$k = \frac{d\varphi}{ds}.$$

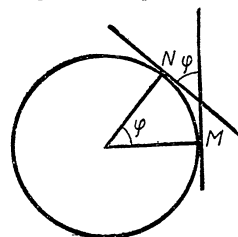


Рис. 222.

В качестве примера покажем, что кривизна окружности радиуса R во всех ее точках постоянна и равна $\frac{1}{R}$ (рис. 222). В самом деле,

$$k_{\widehat{MN}} = \frac{\varphi}{\widehat{MN}} = \frac{\varphi}{R\varphi} = \frac{1}{R},$$

откуда, переходя к пределу при $N \rightarrow M$, имеем $k = \frac{1}{R}$.

По аналогии с окружностью величину, обратную кривизне кривой в данной точке этой кривой (если $k \neq 0$), называют *радиусом кривизны кривой* в этой точке и обозначают R : $R = \frac{1}{k}$.

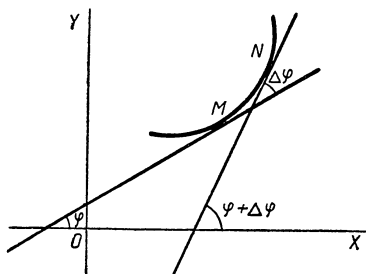


Рис. 223.

Заметим, что иногда полагают $R = \frac{1}{|k|}$ (берется абсолютная величина кривизны) и радиус кривизны рассматривается как положительная величина.

Найдем выражение для кривизны кривой, заданной уравнением $y = f(x)$, где $f(x)$ — дважды дифференцируемая функция. В этом случае (см. рис. 223) $\Delta\varphi$ — угол смежности дуги $\widehat{MN}$, $\operatorname{tg} \varphi = y'$ и, значит, $\varphi = \operatorname{arctg} y'$ (так как рассматривается положительное направление касательной). Следовательно,

$$d\varphi = \frac{y''}{1 + (y')^2} dx.$$

Учитывая, что $ds = \sqrt{1 + (y')^2} dx$ и $k = \frac{d\varphi}{ds}$, получим

$$k = \frac{y''}{[1 + (y')^2]^{\frac{3}{2}}} \quad (1)$$

$$\text{и } R = \frac{[1 + (y')^2]^{\frac{3}{2}}}{y''}.$$

Из формулы для кривизны кривой непосредственно следует, что:

- 1) Знак кривизны кривой $y = f(x)$ в данной точке совпадает со знаком второй производной функции $f(x)$ в этой точке.
- 2) Кривизна прямой $y = kx + b$ во всех ее точках равна нулю ($y'' = 0$), что и естественно, если исходить из обычного представления о прямой. Кривизна кривой в точке перегиба также равна нулю, т. е. в точке перегиба кривая как бы выпрямляется.
- 3) Кривизна кривой в точке экстремума равна y'' .

Пример. Найти кривизну кривой $y = \ln x$ в точке $x = 1$.
Решение.

$$y' = \frac{1}{x}, \quad y'|_{x=1} = 1,$$

$$y'' = -\frac{1}{x^2}, \quad y''|_{x=1} = -1,$$

следовательно, $k = -\frac{1}{\frac{1}{3}} = -\frac{\sqrt{2}}{4}$.

Кривизна кривой, заданной параметрическими уравнениями. Пусть $x = \varphi(t)$, $y = \psi(t)$ — параметрические уравнения кривой, где $\varphi(t)$ и $\psi(t)$ — дважды дифференцируемые функции на некотором отрезке изменения t . Кроме того, предполагается, что эту кривую на рассматриваемом отрезке изменения t можно записать в виде: $y = f(x)$. Используя полученную формулу для кривизны кривой, а также формулы параметрического дифференцирования, запишем:

$$k = \frac{y''}{[1 + (y')^2]^{\frac{3}{2}}} = \frac{\frac{\dot{x}\ddot{y} - \dot{y}\dot{x}}{\dot{x}^3}}{\left[1 + \left(\frac{\dot{y}}{\dot{x}}\right)^2\right]^{\frac{3}{2}}} = \frac{\dot{x}\ddot{y} - \dot{y}\dot{x}}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}},$$

т. е.

$$k = \frac{\dot{x}\ddot{y} - \dot{y}\dot{x}}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}}. \quad (2)$$

Заметим, что, проводя преобразования, вместо $\sqrt{\dot{x}^2} = |\dot{x}|$, мы записали $\dot{x}$, а это означает, что равенство (2) получено в предположении $\dot{x} > 0$ (x возрастает в некоторой окрестности рассматриваемой точки). Равенство (2) справедливо также и в случае, когда $\dot{x} < 0$ (x убывает в некоторой окрестности рассматриваемой точки). Проверить это предоставляем читателю.

§ 26. Соприкасающийся круг. Центр кривизны

Рассмотрим кривые с уравнениями $y = f(x)$ и $y = \varphi(x)$. Будем говорить, что эти кривые имеют в точке x_0 соприкосновение нулевого порядка, если

$$f(x_0) = \varphi(x_0),$$

а

$$f'(x_0) \neq \varphi'(x_0).$$

Геометрически это означает, что эти кривые пересекаются в точке $x = x_0$ (рис. 224). Будем говорить, что кривые $y = f(x)$ и $y = \varphi(x)$ имеют в точке x_0 соприкосновение первого порядка, если

$$f(x_0) = \varphi(x_0),$$

$$f'(x_0) = \varphi'(x_0),$$

а

$$f''(x_0) \neq \varphi''(x_0).$$

Геометрически это означает, что кривые (рис. 224) имеют общую точку $x = x_0$ и имеют общую касательную в этой точке.

Аналогично определяется соприкосновение n -го порядка в точке x_0 . В этом случае

$$f(x_0) = \varphi(x_0),$$

$$f'(x_0) = \varphi'(x_0),$$

$$\dot{f}^{(n)}(x_0) = \dot{\varphi}^{(n)}(x_0),$$

а

$$f^{(n+1)}(x_0) \neq \varphi^{(n+1)}(x_0).$$

(При этом мы предполагали существование в точке x_0 производных тех порядков от функций $f(x)$ и $\varphi(x)$, которые фигурируют в соответствующих определениях.)

Рассмотрим пример. Если функция $f(x)$ $n+1$ раз непрерывно дифференцируема в точке x_0 , то график ее многочлена Тейлора

$$P(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

имеет с графиком функции $f(x)$ в точке x_0 соприкосновение n -го порядка, точнее, соприкосновение порядка не ниже n (порядок соприкосновения будет выше n , если $f^{(n+1)}(x_0) = 0$, так как в этом случае $f^{(n+1)}(x_0) = P^{(n+1)}(x_0) = 0$). Таким образом, в достаточно малой окрестности точки x_0 график многочлена Тейлора n -й степени функции $f(x)$ «наиболее тесно прилегает» к графику самой функции $f(x)$ по сравнению со всеми другими графиками многочленов, степени которых не больше n .

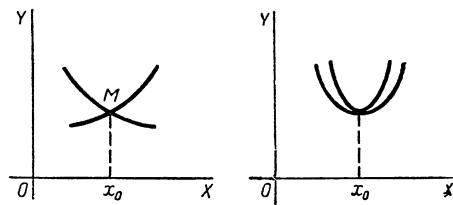


Рис. 224.

Задача 1. Соприкасающаяся прямая. Из всех прямых $y = kx + b$ требуется найти ту, которая имеет порядок соприкосновения с кривой $y = f(x)$ в точке x_0 не ниже первого.

Решение. Для определения двух параметров k и b запишем два условия, которые получаются из условий соприкосновения кривой $y = f(x)$ и прямой $y = kx + b$ в точке $M(x_0, y_0)$:

$$\begin{aligned} f(x_0) &= kx_0 + b, \\ f'(x_0) &= k, \end{aligned}$$

откуда определим:

$$\begin{aligned} k &= f'(x_0) \\ b &= f(x_0) - f'(x_0)x_0. \end{aligned}$$

Уравнение соприкасающейся прямой запишется в виде:

$$y = f'(x_0)x + f(x_0) - f'(x_0)x_0,$$

или

$$y - f(x_0) = f'(x_0)(x - x_0),$$

а это, как известно, уравнение касательной к кривой в точке x_0 . Таким образом, не вертикальная касательная к кривой в данной точке обладает тем свойством, что она имеет наибольший, по сравнению с другими прямыми, порядок соприкосновения с кривой в этой точке.

Задача 2. Соприкасающийся круг¹⁾. Рассмотрим совокупность всех окружностей $(x - \xi)^2 + (y - \eta)^2 = R^2$. Требуется определить параметры ξ , η , R так, чтобы соответствующая окружность с кривой $y = f(x)$ имела в точке $M(x_0, y_0)$ соприкосновение порядка не ниже второго (по сравнению с другими окружностями плоскости).

Для определения трех параметров запишем три условия, которые получаются из условий соприкосновения данной кривой и окружности в точке $M(x_0, y_0)$.

1-е условие. Окружность проходит через точку $M(x_0, y_0)$, т. е.

$$(x_0 - \xi)^2 + (y_0 - \eta)^2 = R^2. \quad (1)$$

Обозначим $y'_0 = f'(x_0)$.

¹⁾ Здесь слово «круг» употребляется в смысле «окружность».

Дифференцируя по x уравнение окружности, получим:

$$2(x - \xi) + 2(y - \eta)y' = 0,$$

или

$$(x - \xi) + (y - \eta)y' = 0. \quad (1')$$

2-е условие. Производные в точке $M(x_0, y_0)$ равны, т. е.

$$(x_0 - \xi) + (y_0 - \eta)y'_0 = 0. \quad (2)$$

Обозначим $y''_0 = f''(x_0)$.

Дифференцируя равенство (1'), получим:

$$1 + (y'')^2 + (y - \eta)y'' = 0.$$

3-е условие. Вторые производные в точке $M(x_0, y_0)$ равны, следовательно,

$$1 + (y''_0)^2 + (y_0 - \eta)y''_0 = 0. \quad (3)$$

Из системы трех уравнений (1), (2), (3) определим значения параметров ξ, η, R :

$$\begin{aligned} (x_0 - \xi)^2 + (y_0 - \eta)^2 &= R^2, \\ (x_0 - \xi) + (y_0 - \eta)y'_0 &= 0, \\ 1 + (y'_0)^2 + (y_0 - \eta)y''_0 &= 0. \end{aligned}$$

Например, из уравнений (2), (3), предполагая, что $y''_0 \neq 0$, найдем:

$$\left. \begin{aligned} y_0 - \eta &= -\frac{1 + (y'_0)^2}{y''_0}, \\ x_0 - \xi &= \frac{1 + (y'_0)^2}{y''_0}y'_0, \end{aligned} \right\} \quad (4)$$

откуда

$$\left. \begin{aligned} \xi &= x_0 - \frac{1 + (y'_0)^2}{y''_0}y'_0, \\ \eta &= y_0 + \frac{1 + (y'_0)^2}{y''_0}. \end{aligned} \right\} \quad (5)$$

Подставляя равенства (4) в первое уравнение, получим:

$$R = \frac{[1 + (y'_0)^2]^{\frac{3}{2}}}{|y''_0|}. \quad (6)$$

Из равенств (5) и (6) определяются координаты центра и радиус окружности, которая имеет наибольший, по сравнению с другими окружностями, порядок соприкосновения с кривой в данной точке $M(x_0, y_0)$.

Окружность $(x - \xi)^2 + (y - \eta)^2 = R^2$, где ξ, η, R определяются из равенств (5) и (6), называется *соприкасающимся кругом кривой $y = f(x)$ в данной точке M* .

Заметим, что соприкасающийся круг кривой в точке M можно было бы определить как предельное положение окружности, проходящей через три точки кривой, при стремлении их к совпадению с точкой M (сравните с определением касательной).

§ 27. Круг кривизны кривой. Эволюта и эвольвента

*Кругом*¹⁾ *кривизны кривой* в ее точке M называется окружность, касающаяся кривой в этой точке и имеющая кривизну, равную кривизне кривой в этой точке.

¹⁾ См. сноску на стр. 257.

Нетрудно видеть, что окружностей, касающихся кривой в точке M , бесчисленное множество, достаточно взять любую окружность, проходящую через точку M , с центром на нормали к кривой в этой точке. Но по определению круга кривизны кривизна кривой в точке M равна кривизне соответствующей окружности, а это означает, что: 1) центр круга кривизны лежит на нормали к кривой в сторону вогнутости кривой, 2) радиус круга кривизны равен радиусу кривизны кривой в точке M :

$$R = \frac{[1 + (y')^2]^{\frac{3}{2}}}{|y''|}.$$

Учитывая сказанное, легко построить (рис. 225) круг кривизны кривой в точке M .

Центр круга кривизны кривой в точке M называется *центром кривизны кривой* в этой точке.

Теорема. Соприкасающийся круг кривой в точке M является кругом кривизны кривой в этой точке.

Действительно:

1) Обе окружности касаются кривой в точке M , следовательно, их центры лежат на нормали к кривой в точке M .

2) Их центры лежат на нормали в сторону вогнутости кривой, что следует из совпадения знаков вторых производных $y_{кр}$ и $y_{окр}$.

3) Радиусы этих окружностей также равны, как это следует из формулы (6) § 26.

На основании этой теоремы получаем, что координаты центра кривизны кривой определяются из равенства (5) § 26, т. е.

$$\left. \begin{aligned} \xi &= x - \frac{1 + (y')^2}{y''} y', \\ \eta &= y + \frac{1 + (y')^2}{y''}. \end{aligned} \right\} \quad (1)$$

Пример. Найти уравнение круга кривизны кривой $y = \ln x$ в точке $x = 1$.
Решение.

$$y = \ln x, \quad y(1) = 0,$$

$$y' = \frac{1}{x}, \quad y'(1) = 1,$$

$$y'' = -\frac{1}{x^2}, \quad y''(1) = -1.$$

Подставляя значения функции и производных при $x = 1$ в формулы для координат центра и радиуса, найдем:

$$\xi = 1 - \frac{2}{-1} = 3, \quad \eta = 0 - 2 = -2, \quad R = 2\sqrt{2}$$

и уравнение круга кривизны запишется в виде:

$$(x - 3)^2 + (y + 2)^2 = 8$$

(рис. 226).

Заметим, что если использовать следующие известные нам равенства:

$$R = \frac{[1 + (y')^2]^{\frac{3}{2}}}{y''}, \quad \sin \alpha = \frac{dy}{ds} = \frac{y'}{\sqrt{1 + (y')^2}},$$

$$\cos \alpha = \frac{dx}{ds} = \frac{1}{\sqrt{1 + (y')^2}},$$

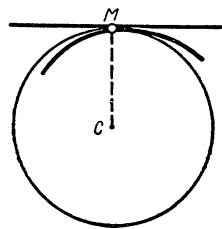


Рис. 225.

то формулы (1) координат центра кривизны можно записать в виде:

$$\left. \begin{aligned} \xi &= x - R \sin \alpha, \\ \eta &= y + R \cos \alpha. \end{aligned} \right\} \quad (2)$$

Заметим также, что равенства (2) можно получить и геометрически (см. рис. 227).

Эволюта плоской кривой. Если точка $M(x, y)$ перемещается по некоторой кривой, то соответствующий ей центр кривизны, вообще говоря,

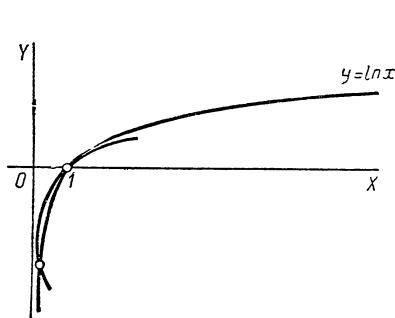


Рис. 226.

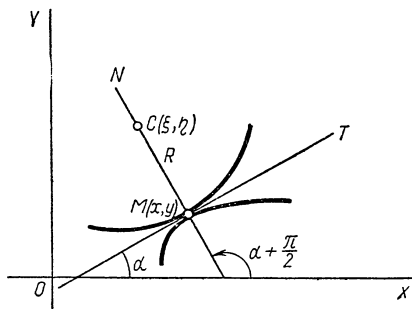


Рис. 227.

тоже перемещается по кривой, которая называется *эволютой* данной кривой. Другими словами, эволютой плоской кривой называется геометрическое место центров кривизны этой кривой. Данная кривая по отношению к своей эволюте называется *эвольвентой*.

Рассмотрим теоремы, характеризующие свойства эволюты.

Теорема 1. Нормаль к данной кривой является касательной к ее эволюте.

Теорема 2. Модуль разности радиусов кривизны в любых точках кривой равен модулю длины соответствующей дуги эволюты.

Например (рис. 228),

$$\begin{aligned} |R_2 - R_1| &= C_1 C_2, \\ |R_3 - R_2| &= C_2 C_3. \end{aligned}$$

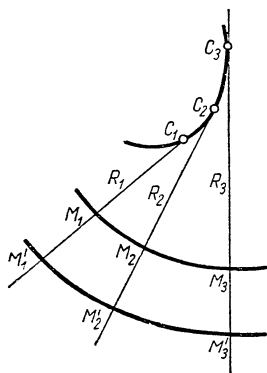


Рис. 228.

Эти свойства эволюты имеют следующие толкования: если на эволюту¹⁾ натянута нить, то эвольвента²⁾ получается, как траектория конца нити при ее сматывании (или наматывании) при условии, что нить находится в натянутом состоянии. Данная эволюта имеет бесчисленное множество эвольвент, причем любые две эвольвенты эквидистантны, т. е. имеют общие нормали и расстояния между ними по нормальям равны.

Как известно, равенства (1) и (2) определяют координаты центра кривизны кривой $y = f(x)$. Если рассматривать x как параметр, то геометрическое место точек с координатами (ξ, η) и есть эволюта рассматриваемой кривой. Таким образом, параметрические уравнения эволюты записываются в виде:

$$\left. \begin{aligned} \xi &= x - \frac{1 + (y')^2}{y''} \cdot y', \\ \eta &= y + \frac{1 + (y')^2}{y''} \end{aligned} \right\} \quad (1)$$

¹⁾ Эволюта (лат.) — кривая, которая развертывается.

²⁾ Эвольвента (лат.) — кривая, которая является разверткой другой кривой.

или в виде

$$\left. \begin{aligned} \xi &= x - R \sin \alpha, \\ \eta &= y + R \cos \alpha. \end{aligned} \right\} \quad (2)$$

При этом параметром в равенствах (1) и (2) может быть как x , так и любой другой параметр, от которого зависят x и y .

Доказательство теоремы 1. Дифференцируя равенства (2), получим¹⁾:

$$\begin{aligned} d\xi &= dx - R \cos \alpha d\alpha - dR \sin \alpha, \\ d\eta &= dy - R \sin \alpha d\alpha + dR \cos \alpha. \end{aligned}$$

Но

$$\begin{aligned} R \cos \alpha d\alpha &= \frac{ds}{d\alpha} \cdot \frac{dx}{ds} d\alpha = dx, \\ R \sin \alpha d\alpha &= \frac{ds}{d\alpha} \cdot \frac{dy}{ds} d\alpha = dy, \end{aligned}$$

потому

$$\left. \begin{aligned} d\xi &= -\sin \alpha dR, \\ d\eta &= \cos \alpha dR. \end{aligned} \right\} \quad (2')$$

Предположим, что R монотонно вырастает или убывает, т. е. $dR \neq 0$, тогда в результате деления равенств (2') имеем:

$$\frac{d\eta}{d\xi} = -\operatorname{ctg} \alpha = -\frac{1}{\operatorname{tg} \alpha}$$

или, что то же,

$$\frac{d\eta}{d\xi} = -\frac{1}{\frac{dy}{dx}}.$$

А это и означает, что касательные к эволюте и эвольвенте в соответствующих точках взаимно перпендикулярны, что и требовалось доказать.

Доказательство теоремы 2. На основании равенств (2') запишем выражение для дифференциала дуги ds эволюты:

$$ds^2 = d\xi^2 + d\eta^2 = dR^2,$$

откуда $\left(\frac{ds}{dR}\right)^2 = 1$ и $\left|\frac{ds}{dR}\right| = 1$.

Если, например, с возрастанием R длина дуги также возрастает, то $\frac{ds}{dR} = 1$ и, значит, $dR = ds$, а следовательно²⁾, $R = s + C$, где C — любое число, т. е. произвольная постоянная. Из соотношения $R = s + C$ ($C = \text{const}$), получаем: $\Delta R = \Delta s$ — приращение радиуса кривизны кривой равно приращению длины дуги эволюты (как по величине, так и по знаку).

Нетрудно заметить, что во всех случаях (при условии, что R изменяется монотонно) абсолютные величины приращений радиуса кривизны и длины дуги равны, что и требовалось доказать.

Заметим еще, что ΔR и Δs имеют одинаковые знаки, если длина дуги эволюты, как функция радиуса кривизны, возрастающая, и имеют противоположные знаки, если длина дуги убывающая функция переменной R .

¹⁾ При этом предполагается, что такое дифференцирование возможно. Например, в случае задания кривой в виде $y = f(x)$ достаточно предположить, что $f(x)$ имеет на рассматриваемом отрезке изменения x непрерывную третью производную.

²⁾ Ясно, что если $R = s + C$ ($C = \text{const}$), то $dR = ds$. Обратно, из соотношения $dR = ds$ следует $R = s + C$, где C — любое число (см. гл. V).

Пример. Написать уравнение эволюты кривой

$$\begin{cases} x = a(\cos t + t \sin t), \\ y = a(\sin t - t \cos t). \end{cases}$$

Решение.

$$\begin{aligned} \begin{cases} \dot{x} = a(-\sin t + \sin t + t \cos t) = at \cos t, \\ \dot{y} = a(\cos t - \cos t + t \sin t) = at \sin t, \end{cases} \\ y' = \frac{\dot{y}}{\dot{x}} = \operatorname{tg} t, \quad 1 + (y')^2 = \sec^2 t, \\ y'' = \frac{d}{dx}(\operatorname{tg} t) = \frac{d}{dt}(\operatorname{tg} t) \cdot \frac{dt}{dx} = \sec^2 t \cdot \frac{1}{x} = \frac{\sec^3 t}{at}. \end{aligned}$$

Уравнения эволюты:

$$\begin{cases} \xi = a(\cos t + t \sin t) - \frac{at \sec^2 t}{\sec^3 t} \cdot \operatorname{tg} t = a \cos t, \\ \eta = a(\sin t - t \cos t) + \frac{at \sec^2 t}{\sec^3 t} = a \sin t, \end{cases}$$

т. е.

$$\begin{cases} \xi = a \cos t, \\ \eta = a \sin t. \end{cases}$$

А это, как известно, уравнения окружности с центром в начале координат радиуса a (исключая параметр t , получим $\xi^2 + \eta^2 = a^2$).

Исходная кривая, следовательно, является *разверткой окружности* (рис. 229).

Заметим, что при решении этой задачи можно было бы сначала получить уравнения эволюты в случае параметрического задания кривой, подставляя в равенство (2)

$$y' = \frac{\dot{y}}{\dot{x}} \text{ и } y'' = \frac{\ddot{y}\dot{x} - \dot{y}\ddot{x}}{\dot{x}^3}, \text{ а потом уже решать эту конкретную задачу.}$$

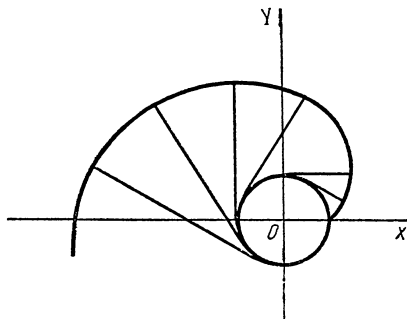


Рис. 229.

§ 28. Приближенное решение уравнений

Рассмотрим уравнение $f(x) = 0$ и пусть ξ — его действительный корень, т. е. $f(\xi) = 0$.

Геометрически равенство $f(\xi) = 0$ означает, что график функции $f(x)$ проходит через точку ξ оси OX . Далее мы будем решать задачу о нахождении с любой наперед заданной точностью приближенного значения корня ξ уравнения $f(x) = 0$, а сначала рассмотрим вопрос о так называемом отделении корней уравнения.

Отделение корней. Будем говорить, что корень ξ уравнения $f(x) = 0$ отделен, если найден отрезок (обозначим его $[a, b]$), в котором, кроме ξ , нет других корней этого уравнения.

Задача отделения корней уравнения $f(x) = 0$ решается просто, если построение графика функции $f(x)$ не представляет большой трудности. Действительно, имея график функции $f(x)$, легко выделить отрезки, в каждом из которых лежит только один корень рассматриваемого уравнения, или, иначе, выделить отрезки, на каждом из которых имеется только одна точка пересечения кривой $y = f(x)$ с осью OX .

Заметим, что если $f(x) = \varphi(x) - \psi(x)$, то отделение корней уравнения $f(x) = 0$ сводится к выделению отрезков, на каждом из которых имеется только одна общая точка кривых $y = \varphi(x)$, $y = \psi(x)$.

Отделить корни уравнения $f(x) = 0$ при условии, что $f(x)$ — дифференцируемая функция, можно и не прибегая к построению графиков. Пусть на концах некоторого отрезка $[a, b]$ функция $f(x)$ имеет значения разных знаков.

Тогда по свойству непрерывных функций эта функция на интервале (a, b) по крайней мере один раз обращается в нуль, т. е. уравнение $f(x) = 0$ имеет по крайней мере один корень.

Если производная $f'(x)$ сохраняет знак на отрезке $[a, b]$, то в силу монотонности функции $f(x)$ уравнение $f(x)=0$ на интервале (a, b) имеет единственный корень.

Приближенное решение уравнений. Пусть корень ξ уравнения $f(x)=0$ отделен, т. е. имеется отрезок $[a, b]$, на котором, кроме ξ , нет других корней этого уравнения. В этом случае числа a и b являются приближенными значениями корня ξ соответственно с недостатком и с избытком.

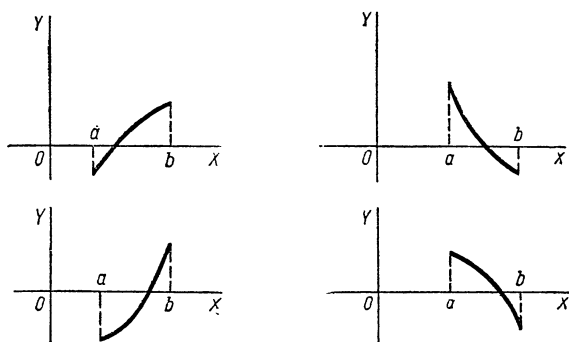


Рис. 230.

Будем искать теперь значение ξ с любой требуемой точностью при следующих допущениях: функция $f(x)$ имеет на отрезке $[a, b]$ непрерывные производные до второго порядка включительно и, кроме того, производные $f'(x)$ и $f''(x)$ сохраняют знаки на этом отрезке.

Из этих условий вытекает, что $f(x)$ — монотонная функция на отрезке $[a, b]$, и следовательно, на концах отрезка она имеет значения разных знаков, а также что кривая $y=f(x)$ выпукла или вогнута (в зависимости от знака $f''(x)$ на рассматриваемом промежутке).

На рис. 230 изображены виды графиков функции в зависимости от знаков $f'(x)$ и $f''(x)$.

Будем определять корень ξ уравнения $f(x)=0$ (предполагается что функция $f(x)$ удовлетворяет перечисленным выше условиям) двумя способами: *хорд* и *касательных*. Смысл этих способов состоит в том, что точка пересечения кривой $y=f(x)$ с осью OX заменяется точкой пересечения с осью OX соответственно хорды (в способе хорд) и касательной (в способе касательных).

Способ хорд. Запишем уравнение хорды AB (рис. 231):

$$\frac{y-f(a)}{f(b)-f(a)} = \frac{x-a}{b-a}.$$

Положив в этом уравнении $y=0$, найдем x_1 — абсциссу точки пересечения хорды AB с осью OX :

$$x_1 = \frac{af(b) - bf(a)}{f(b) - f(a)}.$$

Из условий, налагаемых на функцию $f(x)$, следует, что $a < x_1 < b$. Обозначим через A_1 точку кривой $y=f(x)$, соответствующую $x=x_1$. Рассмотрим хорду A_1B и найдем ее точку пересечения с осью OX :

$$x_2 = \frac{x_1 f(b) - bf(x_1)}{f(b) - f(x_1)},$$

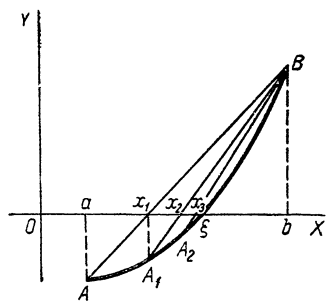


Рис. 231.

при этом $a < x_1 < x_2 < b$. Неограниченно продолжая этот процесс, определим последовательность $\{x_n\}$, при этом

$$x_n = \frac{x_{n-1}f(b) - bf(x_{n-1})}{f(b) - f(x_{n-1})}.$$

Последовательность $\{x_n\}$ — монотонная, ограниченная и, следовательно, сходящаяся. Можно доказать, что $\lim x_n = \xi$. Доказательство этого факта мы не приводим.

Способ касательных. Проведем касательную к кривой $y=f(x)$ в точке B (рис. 232), т. е. в той точке, в которой знаки функций $f(x)$ и $f''(x)$ совпадают (касательная к кривой в точке A может пересечь ось OX за пределами отрезка $[a, b]$), и найдем точку пересечения этой касательной с осью OX .

Уравнение касательной запишется в виде:

$$y - f(b) = f'(b)(x - b).$$

Положив в этом уравнении $y=0$, определим $\bar{x}_1$ — абсциссу точки пересечения касательной с осью OX :

$$-f(b) = f'(b)(\bar{x}_1 - b),$$

$$\bar{x}_1 = b - \frac{f(b)}{f'(b)}.$$

Значению $\bar{x}_1$ соответствует точка кривой $B_1(\bar{x}_1, f(\bar{x}_1))$. Абсцисса точки пересечения касательной к кривой $y=f(x)$ в точке B_1 с осью OX будет

$$\bar{x}_2 = \bar{x}_1 - \frac{f(\bar{x}_1)}{f'(\bar{x}_1)}.$$

Продолжая этот процесс, найдем $\bar{x}_n = \bar{x}_{n-1} - \frac{f(\bar{x}_{n-1})}{f'(\bar{x}_{n-1})}$. Последовательность $\{\bar{x}_n\}$ — монотонная и ограниченная, можно доказать, что ее предел равен ξ .

На практике часто используют одновременно оба способа, одним способом получают приближение искомого корня ξ с недостатком, а другим — приближение с избытком.

Каким именно способом получается приближение корня с недостатком, а каким — с избытком, зависит от вида графика функции $f(x)$.

Если учесть, что каждая из последовательностей $\{x_n\}$ и $\{\bar{x}_n\}$ — монотонная, то нетрудно подсчитать корень с заданной степенью точности, так как совпадающие знаки в приближениях x_n и x_{n+1} (в приближениях $\bar{x}_n$ и $\bar{x}_{n+1}$) являются верными.

Пример. Решим приближенно уравнение $x \ln x = 1$. Сначала запишем эквивалентное ему уравнение $\ln x = \frac{1}{x}$ и построим графики функций

$\ln x$ и $\frac{1}{x}$ (рис. 233). Корень ξ данного уравнения находится в интервале $(2, 3)$. Последнее также легко проверить, так как функция $f(x) = x \ln x - 1$ при $x=2$ отрицательна, $f(2) < 0$, а при $x=3$ — положительна, $f(3) > 0$. Убедимся еще в том, что производные $f'(x)$ и $f''(x)$ сохраняют знаки на отрезке $[2, 3]$. Далее, при $b=3$ имеем $f(3) > 0$,

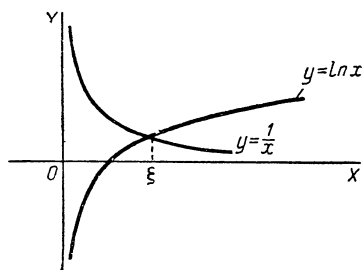


Рис. 233.

$f''(3) > 0$ и по способу касательных получим

$$\bar{x}_1 = 3 - \frac{f(3)}{f'(3)} = 3 - 0,473... \approx 2,527,$$

$$\bar{x}_2 = 2,527 - \frac{f(2,53)}{f'(2,53)} = 2,527 - 0,023... \approx 2,504.$$

Следовательно, $\xi = 2,5$ (при этом мы допускаем погрешность, меньшую чем 0,1).

§ 29. Интерполирование функций

Решение многих вопросов, как теоретического, так и прикладного характера, приводит к следующей задаче: построить полином $P(x)$ степени не выше n , который в $(n+1)$ -й данных точках $x_0, x_1, x_2, \dots, x_n$ принимает соответственно данные значения: $y_0, y_1, y_2, \dots, y_n$. Или, что то же, построить полином степени не выше n , график которого проходит через точки $M_k(x_k, y_k)$ ($k=0, 1, \dots, n$) (рис. 234).

Исторически эта задача возникла из практики составления и употребления математических таблиц. К ней обращаются также, если нужно дать приближенное аналитическое выражение функции (формулу) по известным значениям функции в некоторых точках.

Сначала решим частную задачу: даны точки $x_0, x_1, x_2, \dots, x_n$, построить полином $P_k(x)$ степени n , удовлетворяющий условиям:

$$P_k(x_m) = \begin{cases} 0, & m \neq k, \\ 1, & m = k. \end{cases}$$

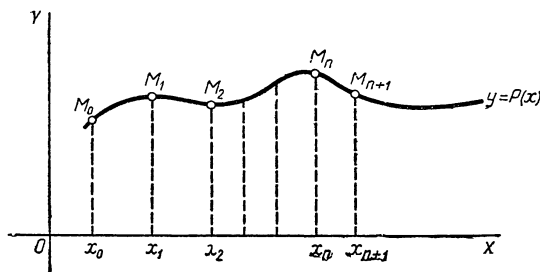


Рис. 234.

График полинома $P_k(x)$ изображен на рис. 235.

Если эта частная задача будет решена, то полином, решающий общую задачу, будет записываться следующим образом:

$$P(x) = y_0 P_0(x) + y_1 P_1(x) + \dots + y_n P_n(x).$$

Действительно, при $x = x_k$ все полиномы $P_m(x)$ ($m \neq k$) равны нулю, а $P_k(x_k) = 1$ и, следовательно, $P(x_k) = y_k$.

Решение частной задачи. Так как полином $P_k(x)$ степени n имеет n нулей в точках $x_0, x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n$, то его можно записать¹⁾ в виде:

$$P_k(x) = A(x - x_0)(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n).$$

Определим коэффициент A , исходя из условия, что $P_k(x_k) = 1$:

$$P_k(x_k) = 1 = A(x_k - x_0)(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n),$$

откуда

$$A = \frac{1}{(x_k - x_0)(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)}.$$

Окончательно полином $P_k(x)$ запишется в виде:

$$P_k(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n)}{(x_k - x_0)(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)}.$$

¹⁾ Разложение полинома на линейные множители рассматривается в § 9 гл. V.

Следовательно, полином $P(x)$, решающий поставленную выше общую задачу, имеет вид:

$$P(x) = \sum_{k=0}^n y_k \frac{(x-x_0)(x-x_1)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{(x_k-x_0)(x_k-x_1)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n)}.$$

Этот полином называется *полиномом Лагранжа*.

Полином $P(x)$ имеет степень не выше n . Это следует из того, что каждое слагаемое полинома имеет степень, равную n , но при сложении коэффициенты при старших степенях n могут взаимно уничтожаться. В частности, если все точки $M_k(x_k, y_k)$, $k=0, 1, \dots, n$, лежат на одной прямой, то $P(x)$ будет полиномом первой или нулевой степени, последнее возможно в случае, когда прямая параллельна оси OX .

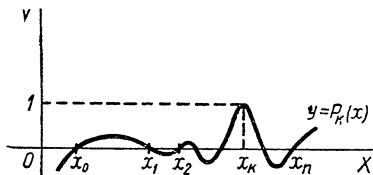


Рис. 235.

Полином $P(x)$ по заданным условиям определяется единственным образом. В самом деле, допустим, что полином $Q(x)$ степени не выше n решает ту же задачу. Тогда разность $P(x) - Q(x) = R(x)$ — полином степени не выше n , который обра-

щается в нуль в $(n+1)$ -й точке x_k ($k=0, 1, \dots, n$). А это возможно лишь при условии, что полином $R(x)$ тождественно равен нулю, т. е. $P(x)$ определяется однозначно.

Другой вид полинома Лагранжа. Рассмотрим полином $A(x)$ степени $n+1$, определяемый равенством

$$A(x) = (x-x_0)(x-x_1)\dots(x-x_n). \quad (1)$$

Легко показать, что

$$A'(x_k) = (x_k-x_0)(x_k-x_1)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n).$$

Учитывая это, можно записать полином Лагранжа в виде:

$$P(x) = \sum_{k=0}^n y_k \frac{A(x)}{(x-x_k) A'(x_k)}. \quad (2)$$

Пример. Построим полином степени $n \leq 2$, принимающий в точках $x_0=1$, $x_1=3$, $x_2=5$ соответственно значения $y_0=2$, $y_1=1$, $y_2=8$.

Решение. По формуле (1) запишем:

$$P(x) = 2 \frac{(x-3)(x-5)}{(1-3)(1-5)} + 1 \frac{(x-1)(x-5)}{(3-1)(3-5)} + 8 \frac{(x-1)(x-3)}{(5-1)(5-3)}.$$

Преобразовав, получим:

$$P(x) = x^2 - \frac{9}{2}x + \frac{11}{2}.$$

Интерполирование функций. Пусть задана некоторая функция $f(x)$ и точки: $x_0, x_1, \dots, x_n$. Полином $P(x)$ степени не выше n , принимающий в данных точках те же значения, что и функция $f(x)$, называется *интерполяционным полиномом*. Точки x_k ($k=0, 1, \dots, n$) в этом случае называются *узлами интерполяции*. По определению интерполяционного полинома имеем:

$$P(x_k) = f(x_k), \quad k=0, 1, 2, \dots, n.$$

Используя выражение полинома Лагранжа в форме (2), получим так называемый *интерполяционный полином Лагранжа*:

$$P(x) = \sum_{k=0}^n f(x_k) \frac{A(x)}{(x-x_k) A'(x_k)}. \quad (3)$$

Вычисление значений этого полинома для значений x , не совпадающих с узлами интерполяции, называется *интерполированием*, или интерполяцией¹⁾.

График интерполяционного полинома пересекает график функции $f(x)$ в узлах интерполяции (рис. 236).

Понятие о конечных разностях. Пусть $f(x)$ — некоторая функция и h — число, отличное от нуля.

Выражение $f(x+h) - f(x)$ называется *конечной разностью первого порядка* (первой конечной разностью) функции $f(x)$ в точке x , она обозначается через $\Delta f(x)$.

$$\Delta f(x) = f(x+h) - f(x).$$

Из определения конечной разности непосредственно следует, что конечная разность суммы функций равна сумме конечных разностей слагаемых и что постоянный множитель можно выносить за знак конечной разности:

$$\begin{aligned}\Delta[f(x) + \varphi(x)] &= \Delta f(x) + \Delta \varphi(x), \\ \Delta[Cf(x)] &= C\Delta f(x).\end{aligned}$$

Конечной разностью второго порядка (второй конечной разностью) функции $f(x)$ в точке x называется выражение $\Delta(\Delta f(x))$, она обозначается через $\Delta^2 f(x)$.

Итак, по определению

$$\Delta^2 f(x) = \Delta(\Delta f(x)).$$

Тогда $\Delta^2 f(x) = \Delta(f(x+h) - f(x)) = \Delta f(x+h) - \Delta f(x) = f(x+2h) - f(x+h) - f(x+h) + f(x)$, откуда следует, что

$$\Delta^2 f(x) = f(x+2h) - 2f(x+h) + f(x).$$

Таким образом, конечная разность второго порядка функции $f(x)$ в точке x , $\Delta^2 f(x)$, выражается через значение функции $f(x)$ в точках x , $x+h$, $x+2h$.

Конечная разность третьего порядка функции $f(x)$ в точке x обозначается через $\Delta^3 f(x)$ и определяется равенством

$$\Delta^3 f(x) = \Delta(\Delta^2 f(x)).$$

Нетрудно получить следующее выражение $\Delta^3 f(x)$ через значения функции $f(x)$ в точках x , $x+h$, $x+2h$, $x+3h$:

$$\Delta^3 f(x) = f(x+3h) - 3f(x+2h) + 3f(x+h) - f(x).$$

Конечная разность n -го порядка функции $f(x)$ в точке x обозначается $\Delta^n f(x)$ и определяется равенством

$$\Delta^n f(x) = \Delta(\Delta^{n-1} f(x)).$$

Можно доказать, что $\Delta^n f(x)$ выражается через значения функции $f(x)$ в точках $x+kh$ ($k=0, 1, \dots, n$) следующим образом:

$$\Delta^n f(x) = \sum_{k=0}^n (-1)^{n-k} C_n^k f(x+kh).$$

Обратно, можно найти выражение значений функции $f(x)$ в точках $x+kh$ через конечные разности в точке x (при этом значение $f(x)$ в точке x будем считать конечной разностью нулевого порядка в этой точке).

¹⁾ Вычисление значений $P(x)$ в точках x , лежащих вне промежутка $[\alpha, \beta]$, где α наименьший, β наибольший из узлов интерполяции, иногда называют экстраполированием, тогда под интерполированием (в узком смысле) понимают случай, когда точка x находится между узлами интерполяции. По-латыни «интер» — между, «экстра» — вне.

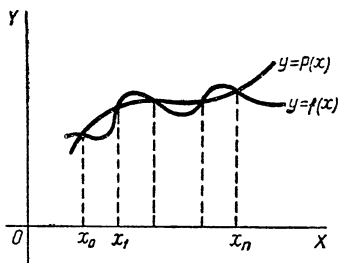


Рис. 236.

Действительно,

$$f(x+h) = f(x) + \Delta f(x), \quad (4)$$

тогда $\Delta f(x+h) = \Delta f(x) + \Delta^2 f(x)$, откуда $f(x+2h) = f(x+h) + \Delta f(x) + \Delta^2 f(x)$ или, учитывая равенство (4),

$$f(x+2h) = f(x) + 2\Delta f(x) + \Delta^2 f(x). \quad (5)$$

Можно доказать, что при любом n

$$f(x+nh) = \sum_{k=0}^n C_n^k \Delta^k f(x). \quad (6)$$

Факториальные полиномы. Рассмотрим систему полиномов, определяемых равенствами:

$$\Phi_0(x) \equiv 1,$$

$$\Phi_n(x) = \frac{1}{n! h^n} (x-x_0)(x-x_0-h) \dots (x-x_0-\overline{n-1}h), \quad n=1, 2, 3, \dots$$

(здесь $\overline{n-1} = n-1$).

Эти полиномы называются *факториальными*.

Полином $\Phi_n(x)$ имеет степень n , его нули образуют арифметическую прогрессию с разностью h .

Докажем следующее свойство факториальных полиномов:

$$\Delta \Phi_{n+1}(x) = \Phi_n(x) \quad (7)$$

(рассматривается конечная разность в любой точке x при приращении, равном h).

Действительно, так как $\Phi_{n+1}(x) = \frac{1}{(n+1)! h^{n+1}} (x-x_0)(x-x_0-h) \dots \dots (x-x_0-nh)$, то

$$\begin{aligned} \Delta \Phi_{n+1}(x) &= \frac{1}{(n+1)! h^{n+1}} [(x+h-x_0)(x-x_0) \dots (x-x_0-\overline{n-1}h) - \\ &\quad - (x-x_0)(x-x_0-h) \dots (x-x_0-\overline{n-1}h)(x-x_0-nh)] = \\ &= \frac{1}{(n+1)! h^{n+1}} (x-x_0)(x-x_0-h) \dots (x-x_0-\overline{n-1}h) \times \\ &\quad \times [x+h-x_0 - (x-x_0-nh)] = \frac{1}{n! h^n} (x-x_0)(x-x_0-h) \dots \\ &\quad \dots (x-x_0-\overline{n-1}h) = \Phi_n(x). \end{aligned}$$

Что и требовалось доказать.

Разложение любого полинома по факториальным полиномам.

Сначала решим общую задачу. Пусть задана какая-то система полиномов $P_0(x)$, $P_1(x)$, ..., $P_n(x)$, степени которых равны их индексам. Докажем, что любой полином $P(x)$ степени n можно разложить по полиномам данной системы, т. е. можно определить числа c_k ($k=0, 1, \dots, n$) так, что при любом x будет выполняться равенство

$$P(x) = c_0 P_0(x) + c_1 P_1(x) + \dots + c_n P_n(x).$$

В самом деле, определим c_n так, чтобы полином $P(x) - c_n P_n(x)$ имел коэффициент при x^n , равный нулю. Этим условием c_n определяется однозначно. Затем подберем c_{n-1} так, чтобы полином $P(x) - c_n P_n(x) - c_{n-1} P_{n-1}(x)$ имел коэффициент при x^{n-1} , равный нулю, и т. д. Наконец, определим c_0 так, чтобы выполнялось равенство

$$P(x) - \sum_{k=0}^n c_k P_k(x) = 0.$$

Из приведенных рассуждений видно, что поставленная задача всегда может быть решена.

Разложим полином $P(x)$ степени n по факториальным полиномам. По доказанному

$$P(x) = c_0 \Phi_0(x) + c_1 \Phi_1(x) + \dots + c_n \Phi_n(x). \quad (8)$$

Определим коэффициенты c_k ($k=0, 1, 2, \dots, n$). Положив $x=x_0$, из равенства (8) найдем $P(x_0)=c_0$. Далее, используя свойство (7) факториальных полиномов, получим:

$$\Delta P(x) = c_1 \Phi_0(x) + c_2 \Phi_1(x) + \dots + c_n \Phi_{n-1}(x),$$

откуда, при $x=x_0$, $\Delta P(x_0)=c_1$.

Рассуждая аналогично, получим:

$$\Delta^2 P(x) = c_2 \Phi_0(x) + c_3 \Phi_1(x) + \dots + c_n \Phi_{n-2}(x),$$

откуда, при $x=x_0$, $\Delta^2 P(x_0)=c_2$ и т. д. Окончательно имеем:

$$\Delta^n P(x) = c_n \Phi_0(x),$$

откуда при $x=x_0$ $\Delta^n P(x_0)=c_n$.

Таким образом, коэффициенты разложения (8) определяются равенствами $c_k = \Delta^k P(x_0)$ ($k=0, 1, 2, \dots, n$) и разложение полинома $P(x)$ по факториальным полиномам записывается в виде:

$$P(x) = P(x_0) + \Delta P(x_0) \Phi_1(x) + \Delta^2 P(x_0) \Phi_2(x) + \dots + \Delta^n P(x_0) \Phi_n(x) \quad (9)$$

или, что то же, в виде:

$$P(x) = P(x_0) + \Delta P(x_0) \frac{x-x_0}{1! h} + \Delta^2 P(x_0) \frac{(x-x_0)(x-x_0-h)}{2! h^2} + \dots + \Delta^n P(x_0) \frac{(x-x_0)(x-x_0-h) \dots (x-x_0-\overline{n-1}h)}{n! h^n}. \quad (10)$$

Пример. Разложить полином $P(x)=x^3$ по факториальным полиномам, взяв $x_0=-2, h=1$.

Решение. Составим таблицу для определения конечных разностей в точке $x_0=-2$.

x	$P(x)=x^3$	$\Delta P(x)$	$\Delta^2 P(x)$	$\Delta^3 P(x)$
-2	-8	7	-6	6
-1	-1	1	0	
0	0	1		
1	1			

$$x^3 = -8\Phi_0(x) + 7\Phi_1(x) - 6\Phi_2(x) + 6\Phi_3(x)$$

или

$$x^3 = -8 + 7 \frac{x+2}{1!} - 6 \frac{(x+2)(x+1)}{2!} + 6 \frac{(x+2)(x+1)x}{3!}.$$

Интерполяционный полином Ньютона. Рассмотрим интерполирование с равноотстоящими узлами (узлы интерполяции $x_k = x_0 + kh$, $k=0, 1, \dots, n$ образуют арифметическую прогрессию с разностью h) и напомним полином $P(x)$ степени не выше n , удовлетворяющий условиям

$$P(x_0 + kh) = f(x_0 + kh), \quad k=0, 1, 2, \dots, n,$$

или, что то же, условиям

$$\Delta^k P(x_0) = \Delta^k f(x_0), \quad k=0, 1, 2, \dots, n.$$

Как уже было указано ранее, такой полином $P(x)$ существует и единственен. Записав его с помощью равенства (10), получим:

$$P(x) = f(x_0) + \Delta f(x_0) \frac{x-x_0}{1! h} + \Delta^2 f(x_0) \frac{(x-x_0)(x-x_0-h)}{2! h^2} + \dots + \Delta^n f(x_0) \frac{(x-x_0)(x-x_0-h) \dots (x-x_0-\overline{n-1}h)}{n! h^n}. \quad (11)$$

Этот полином $P(x)$ называется *интерполяционным полиномом Ньютона*. Так как задача интерполирования с данными узлами интерполяции решается единственным образом, то полином Ньютона от полинома Лагранжа отличается лишь группировкой членов. В формуле Лагранжа каждое слагаемое представляет собой полином степени n , в формуле Ньютона степени полиномов-слагаемых возрастают. Нетрудно видеть, что если добавить новые узлы интерполяции, то в формуле Ньютона добавятся только новые слагаемые, тогда как в формуле Лагранжа придется пересчитывать все слагаемые снова.

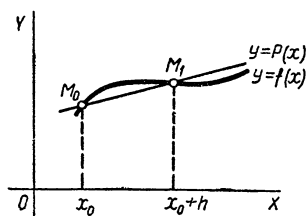


Рис. 237.

На практике чаще применяется интерполирование с равноотстоящими узлами, так как таблицы значений функции составляются обычно для равноотстоящих значений аргумента, поскольку измерение какой-либо величины или, что то же, составление таблицы значений функции производят, как правило, через одинаковые промежутки.

В качестве примеров рассмотрим интерполирование с двумя и тремя узлами интерполяции, т. е. случаи линейной и квадратичной интерполяции.

Линейная интерполяция. Пусть известны значения функции $f(x)$ в двух точках x_0 и x_0+h . Тогда из формулы (11) при $n=1$ получаем

$$P(x) = f(x_0) + \Delta f(x_0) \frac{x-x_0}{h}.$$

Графиком полинома $P(x)$ является прямая (рис. 237).

Будем рассматривать значения x , удовлетворяющие неравенствам $x_0 < x < x_0+h$, и для этих значений x , пренебрегая погрешностью, заменим функцию $f(x)$ полиномом $P(x)$:

$$f(x) = f(x_0) + \Delta f(x_0) \frac{x-x_0}{h},$$

откуда

$$\frac{f(x) - f(x_0)}{f(x_0+h) - f(x_0)} = \frac{x-x_0}{h}.$$

Нетрудно видеть, что это равенство выражает так называемое правило пропорциональных частей, к которому обычно прибегают при работе с таблицами.

Квадратичная интерполяция. Пусть известны значения функции $f(x)$ в трех точках: x_0 , x_0+h , x_0+2h . Из формулы (11) при $n=2$ имеем:

$$P(x) = f(x_0) + \Delta f(x_0) \frac{x-x_0}{1! h} + \Delta^2 f(x_0) \frac{(x-x_0)(x-x_0-h)}{2! h^2}.$$

Если заданные три точки не лежат на одной прямой, то графиком полинома $P(x)$ является парабола второй степени с осью симметрии, параллельной оси OY (рис. 238).

Заменяя функцию $f(x)$ для значений x , удовлетворяющих условию $x_0 \leq x \leq x_0+2h$, интерполяционным полиномом $P(x)$, пренебрегая погрешностью, получим:

$$f(x) = f(x_0) + \Delta f(x_0) \frac{x-x_0}{1! h} + \Delta^2 f(x_0) \frac{(x-x_0)(x-x_0-h)}{2! h^2}.$$

Заметим, что в отличие от линейной интерполяции здесь появляется дополнительное слагаемое $\Delta^2 f(x_0) \frac{(x-x_0)(x-x_0-h)}{2! h^2}$.

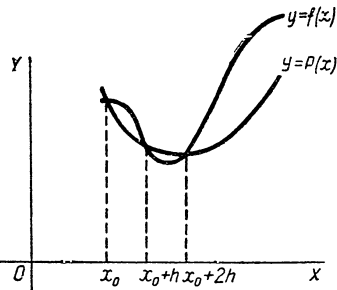


Рис. 238.

Рассмотрим пример.

Вычислим $\lg_{10} 12,7$, используя пятизначные таблицы десятичных логарифмов натуральных чисел. Из таблицы логарифмов находим значение $\lg_{10} x$ при значениях x , равных 12, 13 и 14, и составляем следующую таблицу:

x	$f(x) = \lg_{10} x$	$\Delta f(x)$	$\Delta^2 f(x)$
12	1,07918	0,03476	0,00257
13	1,11394	0,03219	
14	1,14613		

Тогда с помощью линейной интерполяции получим:

$$\lg_{10} 12,7 = 1,07918 + 0,03476 \cdot \frac{0,7}{1} = 1,10351.$$

Если воспользоваться квадратичной интерполяцией, то появится еще одно слагаемое, равное $-\frac{1}{2} 0,00257 \cdot 0,7 \cdot (-0,3) = 0,00027$, и результат квадратичной интерполяции запишется в виде:

$$\lg_{10} 12,7 = 1,10378.$$

Вопрос об оценке точности результата при интерполировании рассматривается дальше.

Другой вид интерполяционной формулы Ньютона. Пусть $\left(\frac{s}{k}\right) = \frac{s(s-1)(s-2)\dots(s-k+1)}{k!}$.

Если положить $s = \frac{x-x_0}{h}$, то формула (11) запишется так:

$$P(x) = f(x_0) + \binom{s}{1} \Delta f(x_0) + \binom{s}{2} \Delta^2 f(x_0) + \dots + \binom{s}{n} \Delta^n f(x_0). \quad (12)$$

Остаточный член интерполяции. Пусть задана функция $f(x)$ и соответствующий интерполяционный полином $P(x)$ с узлами интерполяции $x_k (k=0, 1, \dots, n)$. Разность $f(x) - P(x) = R(x)$ называется *остаточным членом интерполяции*. Оценим величину $R(x)$ для значений x , отличных от узлов интерполяции.

Пусть x — некоторое фиксированное значение, не совпадающее с узлами интерполяции, $\bar{x}$ — переменное. Обозначим

$$A(x) = (x-x_0)(x-x_1)\dots(x-x_n)$$

и

$$K = \frac{R(x)}{A(x)}. \quad (13)$$

Рассмотрим функцию

$$F(\bar{x}) = R(\bar{x}) - KA(\bar{x}) = f(\bar{x}) - P(\bar{x}) - KA(\bar{x}).$$

Функция $F(\bar{x})$ обращается в нуль в узлах интерполяции $x_k (k=0, 1, \dots, n)$ и в точке x , что следует из равенства (13).

Обозначим через α и β соответственно наименьшее и наибольшее из чисел $x_0, x_1, \dots, x_n$ и x . По следствию из теоремы Ролля (гл. IV, § 13) существует точка $\xi, \alpha < \xi < \beta$, такая, что $F^{(n+1)}(\xi) = 0$. Но $F^{(n+1)}(\bar{x}) = f^{(n+1)}(\bar{x}) - K(n+1)!$. Значит, $f^{(n+1)}(\xi) - \frac{R(x)}{A(x)}(n+1)! = 0$, откуда следует, что

$$R(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} A(x), \quad (14)$$

где $\alpha < \xi < \beta$.

Равенство (14) дает выражение остаточного члена интерполяции.

Заметим, что равенство (14) в предельном случае, когда все узлы интерполяции стремятся к совпадению с x_0 , дает выражение остаточного члена формулы Тейлора (гл. IV, § 15). И нетрудно показать, используя несколько раз теорему Ролля, что многочлен Тейлора функции $f(x)$ в точке x_0 является предельным случаем интерполяционного многочлена, когда все узлы интерполяции стремятся к совпадению с x_0 (сделать это предлагается самостоятельно).

Если обозначить через M_{n+1} наибольшее значение функции $|f^{(n+1)}(x)|$ на отрезке $[\alpha, \beta]$, то для остаточного члена интерполяции получаем следующую оценку:

$$|R(x)| \leq \frac{M_{n+1}}{(n+1)!} |A(x)|.$$

В случае равноотстоящих узлов $x_k = x_0 + kh$ ($k=0, 1, \dots, n$)

$$A(x) = (n+1)! h^{n+1} \Phi_{n+1}(x),$$

и

$$|R(x)| \leq M_{n+1} h^{n+1} |\Phi_{n+1}(x)|.$$

Если положить $\frac{x-x_0}{h} = s$, то

$$|R(x)| \leq M_{n+1} h^{n+1} \left| \left(\frac{s}{n+1} \right) \right|.$$

В частности, при линейной интерполяции с узлами x_0, x_0+h для значений x из интервала (x_0, x_0+h) имеем:

$$|R(x)| \leq M_2 h^2 \left| \left(\frac{s}{2} \right) \right|,$$

где M_2 — наибольшее значение $|f''(x)|$ на отрезке $[x_0, x_0+h]$. Но так как $\left| \left(\frac{s}{2} \right) \right| = \left| \frac{s(s-1)}{2} \right|$, где $0 < s < 1$, то следовательно, $\left| \frac{s(s-1)}{2} \right| = \frac{s(1-s)}{2} = \frac{1}{8} - \frac{1}{2} \left(s - \frac{1}{2} \right)^2 \leq \frac{1}{8}$.

Оценка остаточного члена в случае линейной интерполяции имеет вид:

$$|R(x)| \leq \frac{M_2 h^2}{8}.$$

Проводя аналогичные рассуждения в случае квадратичной интерполяции с узлами интерполяции x_0, x_0+h, x_0+2h для значений x из интервала (x_0, x_0+2h) можно получить следующую оценку остаточного члена:

$$|R(x)| \leq \frac{M_3 h^3}{12},$$

где M_3 — наибольшее значение $|f'''(x)|$ на отрезке $[x_0, x_0+2h]$.

ГЛАВА V

НЕОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

Раздел 1

ПЕРВООБРАЗНАЯ ФУНКЦИЯ И НЕОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

Одной из основных задач дифференциального исчисления является отыскание производной заданной функции. Различные вопросы математики, естествознания и техники приводят к решению обратной задачи — восстановлению функции по заданным ее производной или дифференциалу.

Будем называть функцию $F(x)$ *первообразной* (или *примитивной*) для функции $f(x)$ на данном промежутке, если в каждой точке этого промежутка

$$F'(x) = f(x) \quad (1)$$

или, что то же,

$$dF(x) = f(x) dx. \quad (2)$$

Так, например, функция $F(x) = \sin x$ является первообразной на всей оси OX для функции $f(x) = \cos x$, так как для любого x

$$(\sin x)' = \cos x,$$

а функция $\ln|x|$ является первообразной для функции $\frac{1}{x}$ на любом промежутке оси OX , не содержащем точки $x=0$, так как на нем $(\ln|x|)' = \frac{1}{x}$. Отыскание по данной функции $f(x)$ ее первообразной составляет одну из основных задач интегрального исчисления.

Естественно, возникают вопросы: у всякой ли функции существует первообразная и, если существует, то одна или их много?

Ответ на первый вопрос вытекает из доказанной ниже (гл. VI) теоремы: для всякой непрерывной на данном промежутке функции существует первообразная.

В дальнейшем изложении, не оговаривая этого особо, функции, для которых ищутся первообразные, будем считать непрерывными на рассматриваемых промежутках, а данные конкретные элементарные функции будем рассматривать только на промежутках их непрерывности.

Задача отыскания по данной функции $f(x)$ ее первообразной решается неоднозначно: если функция $F(x)$ — первообразная для $f(x)$, то функция $F(x) + C$, где C — любое число, также будет первообразной.

Действительно, если $F(x)$ — первообразная, т. е. $F'(x) = f(x)$, то $[F(x) + C]' = F'(x) = f(x)$ для любого числа C , т. е. $F(x) + C$ — также первообразная.

Например, для функции $f(x) = 3x^2$ функции $F(x) = x^3$; $\Phi(x) = x^3 - \frac{1}{3}$; $\Psi(x) = x^3 + \sqrt{2}$ и вообще $P(x) = x^3 + C$, где C — любое число, будут первообразными.

Таким образом, если функция $f(x)$ имеет одну первообразную $F(x)$, то она имеет их бесчисленное множество.

Покажем теперь, что выражение $F(x) + C$, где C — произвольная постоянная, исчерпывает все первообразные для функции $f(x)$ на данном промежутке.

Лемма. Функция, производная которой на данном промежутке тождественно равна нулю, постоянна на этом промежутке.

Доказательство. Пусть для функции $\varphi(x)$ во всех точках данного промежутка $\varphi'(x) = 0$. Тогда для любых двух его точек x_1 и x_2 по теореме Лагранжа

$$\varphi(x_2) - \varphi(x_1) = \varphi'(\xi)(x_2 - x_1),$$

где ξ — точка, лежащая между точками x_1 и x_2 . Так как $\varphi'(\xi) = 0$, то $\varphi(x_1) = \varphi(x_2)$. Функция, значения которой во всех точках промежутка одинаковы, постоянна, т. е. $\varphi(x) = C$, где C — некоторое число.

Теорема. Если $F(x)$ — первообразная для функции $f(x)$, то любая другая первообразная для $f(x)$ (на том же промежутке) может быть записана в виде $F(x) + C$, где C — некоторое число.

Доказательство. Пусть $\Phi(x)$ — любая первообразная для $f(x)$. Так как $\Phi'(x) = f(x)$, то для всех точек x рассматриваемого промежутка

$$[\Phi(x) - F(x)]' = \Phi'(x) - F'(x) = f(x) - f(x) = 0,$$

а значит, функция $\Phi(x) - F(x)$ постоянна на этом промежутке: $\Phi(x) - F(x) = C$, где C — некоторое число. Отсюда следует, что $\Phi(x) = F(x) + C$.

Итак, если известна одна какая-либо первообразная $F(x)$ функции $f(x)$ (на данном промежутке), то совокупность всех ее первообразных охватывается выражением $F(x) + C$, где C — произвольная постоянная. Другими словами, выражение $F(x) + C$, где $F(x)$ — какая-либо первообразная для $f(x)$, а C — произвольная постоянная, дает все функции, имеющие производной функцию $f(x)$, а дифференциалом $f(x) dx$. Это выражение называется *неопределенным интегралом* от функции $f(x)$ и обозначается символом $\int f(x) dx$ (читается: «интеграл от эф от икс де икс»).

Функция $f(x)$ называется при этом *подинтегральной функцией*, выражение $f(x) dx$ — *подинтегральным выражением*, x — *переменной интегрирования*, символ $\int$ — *знаком интеграла*. Например, $\int \frac{1}{x} dx = \ln|x| + C$, $\int 3x^2 dx = x^3 + C$ и т. д.

Таким образом, по определению

$$\int f(x) dx = F(x) + C, \quad (3)$$

где $F(x)$ — какая-либо первообразная для $f(x)$, а C — произвольная постоянная.

Отыскание первообразной и отыскание неопределенного интеграла от данной функции $f(x)$ называют *интегрированием* этой функции. Интегрирование представляет собой операцию, обратную дифференцированию. Для того чтобы проверить, правильно ли выполнено интегрирование, достаточно продифференцировать результат и получить при этом подинтегральную функцию.

Например, $\int e^{-2x} dx = -\frac{1}{2} e^{-2x} + C$, так как $\left(-\frac{1}{2} e^{-2x} + C\right)' = e^{-2x}$.

§ 1. Простейшие свойства неопределенного интеграла

Из определения неопределенного интеграла [см. (1), (2), (3)] непосредственно следует, что:

1. Производная неопределенного интеграла равна подинтегральной функции:

$$\left(\int f(x) dx\right)' = f(x).$$

2. Дифференциал неопределенного интеграла равен подинтегральному выражению:

$$d\left(\int f(x) dx\right) = f(x) dx.$$

3. Неопределенный интеграл от производной равен самой функции плюс произвольная постоянная C :

$$\int F'(x) dx = F(x) + C.$$

4. Неопределенный интеграл от дифференциала равен дифференцируемой функции плюс произвольная постоянная C :

$$\int dF(x) = F(x) + C.$$

Каждое из свойств 1—4 по-своему выражает тот факт, что интегрирование есть операция, обратная дифференцированию.

5. Постоянный множитель k ($k \neq 0$) можно выносить за знак неопределенного интеграла:

$$\int kf(x) dx = k \int f(x) dx.$$

Доказательство. Пусть $F(x)$ — первообразная для $f(x)$, т. е. $F'(x) = f(x)$. Тогда $kF(x)$ — первообразная для функции $kf(x)$:

$$(kF(x))' = kF'(x) = kf(x).$$

Из определения интеграла следует, что

$$k \int f(x) dx = k[F(x) + C] = kF(x) + C_1 = \int kf(x) dx,$$

где положено $kC = C_1$ (так как $k \neq 0$, то $C_1 = kC$ — тоже произвольная постоянная)¹⁾.

6. Неопределенный интеграл от алгебраической суммы двух функций равен алгебраической сумме интегралов от этих функций (если каждый из них существует):

$$\int [f(x) \pm \varphi(x)] dx = \int f(x) dx \pm \int \varphi(x) dx.$$

Действительно, пусть функции $F(x)$ и $\Phi(x)$ являются первообразными для $f(x)$ и $\varphi(x)$:

$$F'(x) = f(x), \quad \Phi'(x) = \varphi(x).$$

Тогда функции $F(x) \pm \Phi(x)$ будут первообразными для функций $f(x) \pm \varphi(x)$.

Следовательно,

$$\begin{aligned} \int f(x) dx \pm \int \varphi(x) dx &= [F(x) + C_1] \pm [\Phi(x) + C_2] = \\ &= [F(x) \pm \Phi(x)] + [C_1 \pm C_2] = [F(x) \pm \Phi(x)] + C = \\ &= \int [f(x) \pm \varphi(x)] dx, \end{aligned}$$

так как $C = C_1 \pm C_2$ (вместе с C_1 и C_2) — произвольная постоянная.

¹⁾ Если $k=0$, то $k \cdot \int f(x) dx = 0$. В этом случае $k \cdot \int f(x) dx$ будет лишь одной из первообразных для функции $kf(x)$, а $\int kf(x) dx = \int 0 dx = C$, где C — любое число.

Примечание. Свойство 6 легко распространяется на любое число слагаемых.

Таблица основных интегралов. Приведем таблицу так называемых *основных* интегралов. Часть формул этой таблицы непосредственно следует из определения интегрирования как операции, обратной дифференцированию, и таблицы производных. Остальные формулы легко могут быть проверены читателем. Для этого, как уже говорилось выше, достаточно, продифференцировав результат интегрирования, получить подинтегральную функцию.

$$\text{I. } \int x^a dx = \frac{x^{a+1}}{a+1} + C, \quad a - \text{любое действительное число, } a \neq -1,$$

$$\text{II. } \int \frac{dx}{x} = \ln |x| + C,$$

$$\text{III. } \int e^x dx = e^x + C,$$

$$\text{IV. } \int a^x dx = \frac{a^x}{\ln a} + C, \quad a > 0, \quad a \neq 1,$$

$$\text{V. } \int \cos x dx = \sin x + C,$$

$$\text{VI. } \int \sin x dx = -\cos x + C,$$

$$\text{VII. } \int \sec^2 x dx = \operatorname{tg} x + C,$$

$$\text{VIII. } \int \operatorname{cosec}^2 x dx = -\operatorname{ctg} x + C,$$

$$\text{IX. } \int \frac{dx}{x^2 + a^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C, \quad a \neq 0,$$

$$\text{X. } \int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| + C, \quad a \neq 0,$$

$$\text{XI. } \int \frac{dx}{\sqrt{a^2 - x^2}} = \operatorname{arcsin} \frac{x}{a} + C, \quad a > 0,$$

$$\text{XII. } \int \frac{dx}{\sqrt{x^2 - a^2}} = \ln |x + \sqrt{x^2 - a^2}| + C, \quad a \neq 0,$$

$$\text{XII, a. } \int \frac{dx}{\sqrt{x^2 + a^2}} = \ln (x + \sqrt{x^2 + a^2}) + C, \quad a \neq 0.$$

Иногда к основным, или, как их еще называют, *табличным*, интегралам относят также следующие интегралы:

$$\text{XIII. } \int \operatorname{tg} x dx = -\ln |\cos x| + C,$$

$$\text{XIV. } \int \operatorname{ctg} x dx = \ln |\sin x| + C,$$

$$\text{XV. } \int \sec x dx = \ln |\sec x + \operatorname{tg} x| + C = \ln \left| \operatorname{tg} \left(\frac{x}{2} + \frac{\pi}{4} \right) \right| + C,$$

$$\text{XVI. } \int \operatorname{cosec} x dx = \ln |\operatorname{cosec} x - \operatorname{ctg} x| = \ln \left| \operatorname{tg} \frac{x}{2} \right| + C.$$

Проверим, например, справедливость формул IX и XV.

Так как $\left(\frac{1}{a} \operatorname{arctg} \frac{x}{a} + C \right)' = \frac{1}{a} \frac{1}{1 + \frac{x^2}{a^2}} \cdot \frac{1}{a} = \frac{1}{a^2 + x^2}$, то

$$\int \frac{dx}{x^2 + a^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C.$$

Далее, так как $[\ln |\sec x + \operatorname{tg} x| + C]' = \frac{1}{\sec x + \operatorname{tg} x} (\operatorname{tg} x \sec x + \sec^2 x) = \sec x$, то $\int \sec x dx = \ln |\sec x + \operatorname{tg} x| + C$.

Читателю рекомендуется самому проверить остальные формулы.

§ 2. Непосредственное интегрирование

Используя свойства неопределенных интегралов и таблицу основных интегралов, можно интегрировать некоторые элементарные функции (так называемое *непосредственное интегрирование*).

Рассмотрим несколько п р и м е р о в.

$$1. \int x^5 dx = \frac{x^6}{6} + C \text{ (по формуле I).}$$

$$2. \int \frac{dx}{x\sqrt{x}} = \int x^{-\frac{4}{3}} dx = \frac{x^{-\frac{1}{3}}}{-\frac{1}{3}} + C = -3 \frac{1}{\sqrt[3]{x}} + C \text{ (по формуле I).}$$

$$3. \int \frac{dx}{x^2+5} = \frac{1}{\sqrt{5}} \operatorname{arctg} \frac{x}{\sqrt{5}} + C \text{ (по формуле IX).}$$

$$4. \int \frac{dx}{\sqrt{x^2-4}} = \ln |x + \sqrt{x^2-4}| + C \text{ (по формуле XII).}$$

$$5. \int 8 \cos x dx = 8 \int \cos x dx = 8 (\sin x + C_1) = 8 \sin x + C \text{ (здесь } C = 8C_1 \text{).}$$

При решении примера 5 использовано свойство 5 о вынесении постоянного множителя за знак интеграла.

Для того чтобы проинтегрировать сумму функций или, как говорят, взять от нее интеграл, нередко приходится применять свойство 6, заменяя интеграл от суммы функций суммой интегралов от слагаемых.

$$\begin{aligned} 6. \int \left(5x^3 - \frac{8}{x} - \sin x \right) dx &= \int 5x^3 dx - \int \frac{8}{x} dx - \int \sin x dx = 5 \int x^3 dx - \\ &- 8 \int \frac{dx}{x} - \int \sin x dx = \left(\frac{5}{4} x^4 + C_1 \right) - 8 [\ln |x| + C_2] - [-\cos x + C_3] = \\ &= \frac{5}{4} x^4 - 8 \ln |x| + \cos x + C \text{ (по формулам I, II и VI).} \end{aligned}$$

Здесь $C = C_1 - 8C_2 - C_3$ — произвольная постоянная.

П р и м е ч а н и е. При отыскании суммы интегралов обычно сразу пишут одну произвольную постоянную

$$\begin{aligned} 7. \int \frac{dx}{x^2(4+x^2)} &= \frac{1}{4} \int \frac{4dx}{x^2(4+x^2)} = \frac{1}{4} \int \frac{4+x^2-x^2}{x^2(4+x^2)} dx = \\ &= \frac{1}{4} \int \frac{dx}{x^2} - \frac{1}{4} \int \frac{dx}{4+x^2} = -\frac{1}{4} \cdot \frac{1}{x} - \frac{1}{8} \operatorname{arctg} \frac{x}{2} + C \end{aligned}$$

(по свойствам 5 и 6 и формулам I и IX).

Функции, интегрируемые непосредственно, не исчерпывают даже основных элементарных функций, не говоря уже о сложных функциях. Так, например, среди табличных интегралов нет интегралов от $\ln x$, $\arcsin x$, $\operatorname{arctg} x$ и т. д.

Задача интегрирования принципиально труднее задачи дифференцирования. В дифференциальном исчислении имелись конструктивное определение производной¹⁾ и ряд теорем, дающих правила дифференцирования суммы, произведения, частного, сложных и обратных функций. В интегральном исчислении неопределенный интеграл определяется не конструктивно, правил для интегрирования произведения, частного, сложной и обратной функций нет. Имеются лишь отдельные приемы, позволяющие интегрировать отдельные классы функций. Изучением таких приемов мы займемся ниже.

¹⁾ Производная функции $f(x)$ в точке x определялась как предел отношения приращения функции к приращению аргумента при стремлении последнего к нулю.

§ 3. Замена переменной в неопределенном интеграле

Во многих случаях введение новой переменной интегрирования позволяет свести нахождение данного интеграла к нахождению табличного интеграла или интеграла, берущегося тем или иным известным приемом. Такой метод называется *методом подстановки*, а также *методом замены переменной*.

Рассмотрим функцию $x = \varphi(t)$, где $\varphi(t)$ — строго монотонная и непрерывно дифференцируемая функция на некотором промежутке. Если на соответствующем промежутке изменения x функция $f(x)$ интегрируема, то

$$\int f(x) dx = \int f(\varphi(t)) \varphi'(t) dt. \quad (1)$$

Действительно, пусть $F(x)$ — первообразная для $f(x)$:

$$F'(x) = f(x).$$

Тогда функция $F(\varphi(t))$ будет первообразной для функции $f(\varphi(t)) \varphi'(t)$:

$$[F(\varphi(t))]' = F'_x(\varphi(t)) \varphi'(t) = f(\varphi(t)) \varphi'(t)^1.$$

Отсюда следует, что

$$\int f(\varphi(t)) \varphi'(t) dt = F(\varphi(t)) + C = F(x) + C = \int f(x) dx.$$

Формула (1) называется *формулой замены переменной* в неопределенном интеграле.

Заменяя по формуле (1) переменную интегрирования в интеграле $\int f(x) dx$ некоторой функцией t : $x = \varphi(t)$, мы сведем нахождение этого интеграла к нахождению интеграла $\int f(\varphi(t)) \varphi'(t) dt$. Формально замена переменной выполняется так, как будто подинтегральное выражение $f(x) dx$ есть произведение функции $f(x)$ на дифференциал x :

$$f(x) = f(\varphi(t)), \quad dx = d\varphi(t) = \varphi'(t) dt,$$

т. е. $f(x) dx = f(\varphi(t)) \varphi'(t) dt$.

После нахождения интеграла $\int f(\varphi(t)) \varphi'(t) dt$ надо вместо t подставить его выражение через x^2 .

Пусть, например, требуется найти интеграл $\int \frac{x^3}{(x-1)^2} dx$. Обозначим $x = t+1$, тогда $dx = dt$.

$$\begin{aligned} \int \frac{x^3}{(x-1)^2} dx &= \int \frac{(t+1)^3}{t^2} dt = \int \frac{t^3 + 3t^2 + 3t + 1}{t^2} dt = \int \left(t + 3 + \frac{3}{t} + \frac{1}{t^2} \right) dt = \\ &= \frac{1}{2} t^2 + 3t + 3 \ln |t| - \frac{1}{t} + C = \frac{1}{2} (x-1)^2 + 3(x-1) + 3 \ln |x-1| - \frac{1}{x-1} + C. \end{aligned}$$

Замена переменной в неопределенном интеграле чаще производится по формуле (1), прочитанной справа налево:

$$\int f(\varphi(x)) \varphi'(x) dx = \int f(t) dt, \quad (2)$$

где $t = \varphi(x)$.

При этом функция $\varphi(x)$, взятая за новую переменную t , должна быть такова, чтобы обратная к ней функция $x = \psi(t)$ удовлетворяла вышесказанным выше условиям.

¹⁾ Здесь $F'_x(\varphi(t))$ означает производную функции $F(\varphi(t))$ по промежуточному аргументу $x = \varphi(t)$.

²⁾ Существование однозначной обратной функции $t = \psi(x)$ обеспечивается строгой монотонностью функции $x = \varphi(t)$.

Рассмотрим несколько примеров.

1. $\int \frac{x^4 dx}{x^5 + 7}$. Обозначим $x^5 + 7 = t$, тогда $5x^4 dx = dt$ и, следовательно,

$$\int \frac{x^4 dx}{x^5 + 7} = \frac{1}{5} \int \frac{5x^4 dx}{x^5 + 7} = \frac{1}{5} \int \frac{dt}{t} = \frac{1}{5} \ln |t| + C = \frac{1}{5} \ln |x^5 + 7| + C.$$

2. $\int \frac{e^x dx}{\sqrt{4 - e^{2x}}}$. Обозначим $e^x = t$, тогда $e^x dx = dt$ и, значит,

$$\int \frac{e^x dx}{\sqrt{4 - e^{2x}}} = \int \frac{dt}{\sqrt{4 - t^2}} = \arcsin \frac{t}{2} + C = \arcsin \frac{e^x}{2} + C.$$

3. $\int \cos ax dx$, $a \neq 0$. Применим подстановку $ax = t$, тогда $adx = dt$, $\int \cos ax dx = \frac{1}{a} \int \cos ax \cdot adx = \frac{1}{a} \int \cos t dt = \frac{1}{a} \sin t + C = \frac{1}{a} \sin ax + C$.

Аналогично берутся интегралы от e^{ax} , $\sin ax$, $\lg ax$, $\operatorname{ctg} ax$ и т. д.

В приведенных выше примерах метод замены переменной быстро привел к цели. Однако удачный выбор новой переменной обычно представляет известные трудности. Для их успешного преодоления необходимо хорошо владеть техникой дифференцирования, уметь «прикидывать», что даст та или иная подстановка, и твердо знать табличные интегралы.

4. Интегрирование по частям

Пусть функции $u = \varphi(x)$ и $v = \psi(x)$ непрерывно дифференцируемы на некотором промежутке.

Интегрируя тождество

$$[\varphi(x) \psi(x)]' = \varphi'(x) \psi(x) + \varphi(x) \psi'(x),$$

получим:

$$\int [\varphi(x) \psi(x)]' dx = \int \varphi'(x) \psi(x) dx + \int \varphi(x) \psi'(x) dx,$$

т. е.

$$\varphi(x) \psi(x) = \int \varphi'(x) \psi(x) dx + \int \varphi(x) \psi'(x) dx,$$

откуда следует, что

$$\int \varphi(x) \psi'(x) dx = \varphi(x) \psi(x) - \int \psi(x) \varphi'(x) dx. \quad (1)$$

Формулу (1) обычно записывают в виде:

$$\int u dv = uv - \int v du. \quad (1')$$

Эта формула называется *формулой интегрирования по частям*. Правильнее было бы назвать ее формулой частичного интегрирования. При известных u и v она сводит нахождение интеграла от $u dv$ после частичного интегрирования к нахождению интеграла от $v du$. Иногда удается функции u и v выбрать так, что новый интеграл либо сам является табличным, либо сводится к табличным интегралам уже известными методами.

Рассмотрим ряд примеров.

1. $\int x e^x dx$. Полагаем $x = u$, $e^x dx = dv$. Тогда $dx = du$, $e^x = v$ и, значит, по формуле (1)

$$\int x e^x dx = x e^x - \int e^x dx.$$

Интеграл в правой части равенства табличный:

$$\int e^x dx = e^x + C.$$

Окончательно получаем:

$$\int x e^x dx = x e^x - e^x + C_1 \quad (C_1 = -C).$$

Примечание 1. При отыскании v по известному dv достаточно найти одну какую-либо первообразную. Нетрудно видеть, что прибавление к ней произвольной постоянной не меняет результата интегрирования, усложняя выкладки.

Примечание 2. Если в примере 1 положить $e^x = u$, $x dx = dv$, то $e^x dx = du$, $\frac{x^2}{2} = v$ и, значит, по формуле (1)

$$\int x e^x dx = \frac{x^2}{2} e^x - \frac{1}{2} \int x^2 e^x dx.$$

Этот выбор u и dv неудачен, так как новый интеграл «сложнее» данного. В общем случае нельзя указать рецепта для выбора в подинтегральном выражении u и dv . Рекомендуется мысленно «прикидывать», что дадут те или иные обозначения.

2. $\int x^3 \ln |x| dx$. Здесь удобнее положить $u = \ln |x|$, $dv = x^3 dx$. Тогда $du = \frac{dx}{x}$, $v = \frac{x^4}{4}$ и, значит,

$$\int x^3 \ln |x| dx = \frac{x^4}{4} \ln |x| - \frac{1}{4} \int x^3 dx = \frac{x^4}{4} \ln |x| - \frac{x^4}{16} + C.$$

Иногда для получения окончательного результата приходится несколько раз последовательно применять формулу интегрирования по частям

3. $\int x^2 \sin x dx$. Пусть $x^2 = u$, $\sin x dx = dv$, тогда $2x dx = du$, $-\cos x = v$, т. е.

$$\int x^2 \sin x dx = -x^2 \cos x + 2 \int x \cos x dx.$$

Для нахождения интеграла $\int x \cos x dx$ положим $x = u$, $\cos x dx = dv$, тогда $dx = du$, $\sin x = v$,

$$\int x \cos x dx = x \sin x - \int \sin x dx = x \sin x + \cos x + C.$$

Окончательно получаем:

$$\int x^2 \sin x dx = -x^2 \cos x + 2x \sin x + 2 \cos x + C_1,$$

где $C_1 = 2C$.

Метод интегрирования по частям имеет более ограниченную область применения, чем метод замены переменной. Существуют, однако, классы функций, интегрируемых по частям. К ним относятся, например, функции вида $P(x) e^{ax}$, $P(x) \sin ax$, $P(x) \cos ax$, где a — любое действительное число, а $P(x)$ — любой многочлен.

Интегралы от этих функций берутся методом интегрирования по частям, если положить $P(x) = u$, а оставшиеся «множители» взять за dv . После однократного интегрирования по частям получатся интегралы того же вида, но степень многочлена $P(x)$ понизится на единицу. После n -кратного применения формулы (1) останутся интегралы $\int e^{ax} dx$, $\int \sin ax dx$ и $\int \cos ax dx$, сводящиеся к табличным.

4. $\int (x^2 - 2x) e^{-x} dx$. Пусть $x^2 - 2x = u$, $e^{-x} dx = dv$, тогда $2(x - 1) dx = du$, $-e^{-x} = v$, т. е. $\int (x^2 - 2x) e^{-x} dx = -(x^2 - 2x) e^{-x} + 2 \int (x - 1) e^{-x} dx$. Полагаем $x - 1 = u$, $e^{-x} dx = dv$, тогда $dx = du$, $-e^{-x} = v$,

$$\int (x - 1) e^{-x} dx = -(x - 1) e^{-x} + \int e^{-x} dx = -(x - 1) e^{-x} - e^{-x} + C.$$

Окончательно получим:

$$\begin{aligned} \int (x^2 - 2x) e^{-x} dx &= (-x^2 + 2x) e^{-x} + (-2x + 2) e^{-x} - \\ &- 2e^{-x} + C_1 = -x^2 e^{-x} + C_1. \end{aligned}$$

Интегралы вида $\int P(x) \ln |ax| dx$, где a — любое действительное число ($a \neq 0$), а $P(x)$ — любой многочлен, также берутся методом интегрирования по частям, если положить

$$\ln |ax| = u, \quad P(x) dx = dv.$$

В конце данной главы будут дополнительно указаны некоторые классы функций, интегрируемых методом интегрирования по частям с последующим применением других методов и приемов.

Нахождение некоторых интегралов методом интегрирования по частям сводится к решению алгебраического уравнения, в котором неизвестным является искомый интеграл.

Пример. Найти $\int e^{ax} \sin \beta x dx$.

Пусть $e^{ax} = u$, $\sin \beta x dx = dv$. Тогда $ae^{ax} dx = du$, $-\frac{1}{\beta} \cos \beta x = v$, т. е.

$$\int e^{ax} \sin \beta x dx = -\frac{1}{\beta} e^{ax} \cos \beta x + \frac{\alpha}{\beta} \int e^{ax} \cos \beta x dx.$$

Полагаем теперь $e^{ax} = u$, $\cos \beta x dx = dv$. Тогда $ae^{ax} dx = du$, $\frac{1}{\beta} \sin \beta x = v$, т. е.

$$\int e^{ax} \cos \beta x dx = \frac{1}{\beta} e^{ax} \sin \beta x - \frac{\alpha}{\beta} \int e^{ax} \sin \beta x dx.$$

Таким образом,

$$\int e^{ax} \sin \beta x dx = -\frac{1}{\beta} e^{ax} \cos \beta x + \frac{\alpha}{\beta^2} e^{ax} \sin \beta x - \frac{\alpha^2}{\beta^2} \int e^{ax} \sin \beta x dx. \quad (2)$$

Примечание. Так как неопределенный интеграл равен какой-либо первообразной подинтегральной функции плюс произвольная постоянная, то один неопределенный интеграл считается равным выражению, содержащему другой интеграл, если разность между ними постоянна.

Приводя подобные члены в равенстве (2) и учитывая сделанное выше примечание, получим:

$$\begin{aligned} \left(1 + \frac{\alpha^2}{\beta^2}\right) \int e^{ax} \sin \beta x dx &= \frac{-\beta \cos \beta x + \alpha \sin \beta x}{\beta^2} e^{ax} + C_1, \\ \int e^{ax} \sin \beta x dx &= \frac{\alpha \sin \beta x - \beta \cos \beta x}{\alpha^2 + \beta^2} e^{ax} + C, \end{aligned}$$

где $C = \frac{C_1 \beta^2}{\alpha^2 + \beta^2}$ — произвольная постоянная.

Раздел 2 КОМПЛЕКСНЫЕ ЧИСЛА

§ 5. Комплексные числа и действия над ними

В данном параграфе будут кратко изложены действия над комплексными числами, т. е. вопросы, в основном изучаемые в курсе элементарной математики.

Комплексным числом z называют символ $x + yi$, где x и y — действительные числа. При этом x называется *действительной частью* комплексного числа, y — *мнимой частью*, i — *мнимой единицей*.

Комплексные числа $z_1 = x_1 + y_1 i$ и $z_2 = x_2 + y_2 i$ называются *равными*, если соответственно равны их действительные и мнимые части:

$$x_1 = x_2, \quad y_1 = y_2. \quad (1)$$

В частности, комплексное число $z = x + yi = 0 + i0$ тогда, и только тогда, когда $x = 0$, $y = 0$.

Если $y=0$, то комплексное число имеет вид $z=x+0i$. Будем его сокращенно записывать $z=x$ и называть *действительным* числом. Если $x=0$ и $y=0$, то комплексное число $z=0+i0$ сокращенно записывается в виде $z=0$ и называется *нулем*.

Если $x=0$, $y \neq 0$, то комплексное число имеет вид $z=0+yi$ или, короче, $z=yi$. Его называют *чисто мнимым* числом. В частности, если $x=0$, $y=1$, получаем комплексное число $0+1i=i$ — мнимую единицу. Всякое число $z=x+iy$, где $y \neq 0$, называют *мнимым* числом.

Два комплексных числа $x+yi$ и $x-yi$ называются *комплексно-сопряженными*. Если $z=x+yi$, то сопряженное число $x-yi$ принято обозначать $\bar{z}$.

Суммой комплексных чисел $z_1=x_1+y_1i$ и $z_2=x_2+y_2i$ называют комплексное число $(x_1+x_2)+(y_1+y_2)i$. Его обозначают z_1+z_2 . Таким образом,

$$z_1+z_2=(x_1+x_2)+(y_1+y_2)i. \quad (2)$$

При сложении комплексных чисел складываются их действительные и мнимые части.

Комплексное число $z_3=x_3+y_3i$ называется *разностью* двух комплексных чисел $z_1=x_1+y_1i$ и $z_2=x_2+y_2i$, если $z_3+z_2=z_1$. Разность комплексных чисел z_1 и z_2 обозначается z_1-z_2 . Из определения следует, что

$$z_1-z_2=(x_1-x_2)+(y_1-y_2)i. \quad (3)$$

При вычитании из действительной и мнимой части уменьшаемого соответственно вычитается действительная и мнимая часть вычитаемого.

Умножение двух комплексных чисел $z_1=x_1+y_1i$ и $z_2=x_2+y_2i$ вводится равенством

$$z_1 \cdot z_2=(x_1x_2-y_1y_2)+(x_1y_2+x_2y_1)i. \quad (4)$$

Из равенства (4) следует, что

$$i^2=i \cdot i=(0+1i)(0+1i)=-1+0i=-1. \quad (5)$$

Нетрудно заметить, что с учетом того, что $i^2=-1$, равенство (4) может быть формально получено по обычному правилу умножения двучлена на двучлен.

Действительно, $(x_1+y_1i) \cdot (x_2+y_2i)=x_1x_2+y_1y_2i^2+x_1y_2i+x_2y_1i=$
 $= (x_1x_2-y_1y_2)+(x_1y_2+x_2y_1)i$.

Если при умножении двух комплексных чисел z_1 и z_2 получится некоторое число $A+Bi$, то при умножении сопряженных им чисел $\bar{z}_1$ и $\bar{z}_2$ получится число сопряженное, т. е. $A-Bi$. Для доказательства умножим сначала число $z_1=x_1+y_1i$ на $z_2=x_2+y_2i$, а затем число $\bar{z}_1=x_1-y_1i$ на $\bar{z}_2=x_2-y_2i$:

$$z_1z_2=(x_1+y_1i)(x_2+y_2i)=(x_1x_2-y_1y_2)+(x_1y_2+x_2y_1)i=A+Bi,$$

$$\bar{z}_1\bar{z}_2=(x_1-y_1i)(x_2-y_2i)=(x_1x_2-y_1y_2)-(x_1y_2+x_2y_1)i=A-Bi.$$

Деление вводится как действие, обратное умножению. Частным от деления числа $z_1=x_1+y_1i$ на число $z_2=x_2+y_2i$ ($z_2 \neq 0$) называют число $z_3=x_3+y_3i$, такое, что $z_3 \cdot z_2=z_1$, т. е.

$$x_1+y_1i=(x_3+y_3i)(x_2+y_2i). \quad (6)$$

Отсюда, на основании равенства (4), получаем:

$$x_1=x_2x_3-y_2y_3; \quad y_1=x_2y_3+x_3y_2. \quad (7)$$

Решая систему (7) относительно x_3 и y_3 , находим:

$$x_3=\frac{x_1x_2+y_1y_2}{x_2^2+y_2^2}; \quad y_3=\frac{x_2y_1-x_1y_2}{x_2^2+y_2^2} \quad (8)$$

(где $x_2^2+y_2^2 \neq 0$, так как по условию $x_2+y_2i \neq 0$).

Таким образом,

$$\frac{x_1 + y_1 i}{x_2 + y_2 i} = \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + i \frac{x_2 y_1 - x_1 y_2}{x_2^2 + y_2^2}. \quad (9)$$

Нетрудно заметить, что равенство (9) может быть получено путем умножения числителя и знаменателя дроби $\frac{x_1 + y_1 i}{x_2 + y_2 i}$ на число, комплексно-сопряженное знаменателю.

Возведение комплексного числа z в натуральную степень n рассматривается как частный случай умножения комплексных чисел:

$$z^n = \underbrace{z \cdot z \cdot \dots \cdot z}_{n \text{ раз}}$$

При этом, на основании равенства (5), получаем:

$$i^2 = -1; \quad i^3 = i^2 \cdot i = -i; \quad i^4 = i^2 \cdot i^2 = 1,$$

и вообще $i^{4n} = 1$, $i^{4n+1} = i$, $i^{4n+2} = -1$, $i^{4n+3} = -i$, где n — натуральное число.

Примечание 1. Из данных определений следует, что вышеуказанные действия над комплексными числами производятся так же, как и над действительными двучленами, если при этом считать, что $i^2 = -1$.

Примечание 2. Если над несколькими комплексными числами совершается ряд алгебраических операций (сложение, вычитание, умножение, деление, возведение в целую положительную степень), в результате которых получается некоторое число $A + Bi$, то при совершении тех же операций над числами, сопряженными данным, получится число $A - Bi$, сопряженное числу $A + Bi$.

Это утверждение было доказано в случае умножения комплексных чисел. Таким же образом может быть оно доказано и для остальных операций.

Например, если $\frac{(z_1 + z_2) z_3}{z_4 - z_5} = A + Bi$, то $\frac{(\bar{z}_1 + \bar{z}_2) \bar{z}_3}{\bar{z}_4 - \bar{z}_5} = A - Bi$.

Комплексные числа можно рассматривать как расширение множества действительных чисел. В самом деле, алгебраические операции над комплексными числами введены так, что совокупность всех «действительных» комплексных чисел (т. е. чисел вида $z = x + i0$ или, короче, $z = x$) с указанными операциями над ними совпадает с совокупностью действительных чисел и известными действиями над этими числами.

§ 6. Тригонометрическая форма комплексного числа

Геометрическая интерпретация комплексного числа. Комплексное число $z = x + yi$ геометрически изображается точкой с координатами x, y на плоскости $ХОУ$ (рис. 239). Плоскость, служащая для изображения комплексных чисел, называется *комплексной плоскостью*. Так как каждое комплексное число $z = x + yi$ определяется его действительной частью x и мнимой частью y единственным образом, то каждому комплексному числу будет соответствовать единственная точка в комплексной плоскости. Очевидно, что справедливо и обратное положение, а именно, что каждая точка плоскости $ХОУ$ изображает единственное комплексное число. Следовательно, между комплексными числами и точками комплексной плоскости существует взаимно однозначное соответствие.

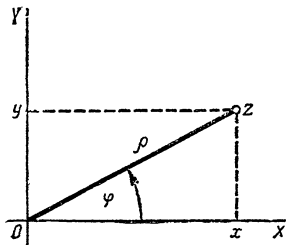


Рис. 239.

Если $z_1 = x_1 + y_1 i$ — действительное число (т. е. $y_1 = 0$, $z_1 = x_1$), то соответствующая точка лежит на действительной оси. Поэтому ось абсцисс называется *действительной осью*. Если комплексное число $z_2 = x_2 + y_2 i$ — чисто мнимое (т. е. $x_2 = 0$, $z_2 = y_2 i$), то соответствующая точка лежит на оси ординат. Поэтому ось ординат называют *мнимой осью*.

Если каждой точке z комплексной плоскости отнести вектор $\vec{Oz}$, то появится возможность представлять комплексные числа векторами.

Тригонометрическая форма комплексного числа. Выберем на плоскости $ХОУ$ полярную систему координат (рис. 239) так, чтобы полюс совпал с началом координат, а полярная ось пошла бы по положительному направлению действительной оси. Обозначим полярный радиус точки $z = x + yi$ через ρ , а полярный угол через φ . Полярный радиус ρ называется *модулем* комплексного числа и обозначается $|z|$. Полярный угол φ называется *аргументом* комплексного числа и обозначается $\arg z$, если берется главное значение угла ($0 \leq \arg z < 2\pi$), и $\text{Arg } z$, если берется общее значение угла. Таким образом,

$$\rho = |z|, \quad \text{Arg } z = \varphi + 2\pi k, \quad (1)$$

где k — произвольное целое число, а φ — любое из значений аргумента z . Так как $x = \rho \cos \varphi$, а $y = \rho \sin \varphi$, то

$$\begin{aligned} z &= x + yi = \rho \cos \varphi + i \rho \sin \varphi, \\ z &= \rho (\cos \varphi + i \sin \varphi). \end{aligned} \quad (2)$$

Выражение (2) называется *тригонометрической формой* записи комплексного числа. Очевидно, что

$$|z| = \sqrt{x^2 + y^2}, \quad \tg \varphi = \frac{y}{x}. \quad (3)$$

Например, $2 + 2\sqrt{3}i = 4 \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right)$.

Геометрическая интерпретация сложения комплексных чисел. Пусть в плоскости комплексной переменной даются два числа z_1 и z_2 (рис. 240).

Проведя радиусы-векторы точек z_1 и z_2 , получим два вектора $\vec{Oz}_1$ и $\vec{Oz}_2$, которые соответствуют комплексным числам z_1 и z_2 . При сложении комплексных чисел складываются их действительные и мнимые части, а при сложении векторов складываются соответствующие координаты. Это позволяет сложение комплексных чисел представлять в виде сложения векторов. Вектор $\vec{Oz}_3$, являющийся диагональю параллелограмма, построенного на векторах $\vec{Oz}_1$ и $\vec{Oz}_2$, и представляет комплексное число: $z_3 = z_1 + z_2$.

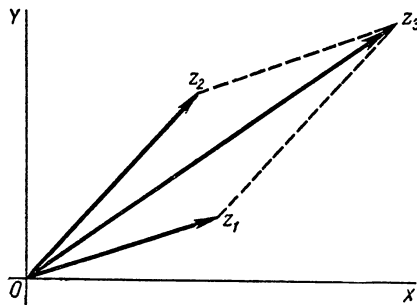


Рис. 240.

§ 7. Действия над комплексными числами в тригонометрической форме

Умножение комплексных чисел. Возьмем два комплексных числа в тригонометрической форме:

$$z_1 = \rho_1 (\cos \varphi_1 + i \sin \varphi_1) \quad \text{и} \quad z_2 = \rho_2 (\cos \varphi_2 + i \sin \varphi_2).$$

Перемножая их, получим:

$$\begin{aligned} z_1 z_2 &= \rho_1 \rho_2 (\cos \varphi_1 + i \sin \varphi_1) \cdot (\cos \varphi_2 + i \sin \varphi_2) = \\ &= \rho_1 \rho_2 (\cos \varphi_1 \cos \varphi_2 + i \sin \varphi_1 \cos \varphi_2 + i \cos \varphi_1 \sin \varphi_2 - \sin \varphi_1 \sin \varphi_2) = \\ &= \rho_1 \rho_2 [(\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) + i (\sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2)] = \\ &= \rho_1 \rho_2 [\cos (\varphi_1 + \varphi_2) + i \sin (\varphi_1 + \varphi_2)], \end{aligned}$$

т. е.

$$z_1 z_2 = \rho_1 \rho_2 [\cos (\varphi_1 + \varphi_2) + i \sin (\varphi_1 + \varphi_2)]. \quad (1)$$

Таким образом, при умножении комплексных чисел модули их перемножаются:

$$|z_1 z_2| = \rho_1 \rho_2 = |z_1| \cdot |z_2|, \quad (2)$$

а аргументы складываются

$$\text{Arg} (z_1 z_2) = \text{Arg} z_1 + \text{Arg} z_2. \quad (3)$$

Геометрическая интерпретация умножения комплексных чисел. Пусть требуется умножить комплексное число $z_1 = \rho_1 (\cos \varphi_1 + i \sin \varphi_1)$ на комплексное число $z_2 = \rho_2 (\cos \varphi_2 + i \sin \varphi_2)$.

Комплексное число z_1 представляется вектором $\vec{Oz}_1$, а комплексное число z_2 представляется вектором $\vec{Oz}_2$. Чтобы построить вектор, изображающий комплексное число

$$z_3 = z_1 \cdot z_2 = \rho_1 \rho_2 [\cos (\varphi_1 + \varphi_2) + i \sin (\varphi_1 + \varphi_2)],$$

надо вектор $\vec{Oz}_1$ повернуть на угол φ_2 против часовой стрелки, если $\varphi_2 > 0$, и по часовой стрелке, если $\varphi_2 < 0$, и «увеличить» его длину в ρ_2 раз (рис. 241). Так как

$i = 1 \cdot \left(\cos \frac{\pi}{2} + i \sin \frac{\pi}{2} \right)$, то умножение любого числа z на i можно рассматривать как операцию поворота вектора, изображающего число z на угол $\frac{\pi}{2}$ против движения часовой стрелки.

Деление комплексных чисел. Возьмем два комплексных числа $z_1 = \rho_1 (\cos \varphi_1 + i \sin \varphi_1)$ и $z_2 = \rho_2 (\cos \varphi_2 + i \sin \varphi_2)$ и разделим z_1 на z_2 .

Так как деление комплексных чисел рассматривается как действие, обратное умножению, то

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2} [\cos (\varphi_1 - \varphi_2) + i \sin (\varphi_1 - \varphi_2)]. \quad (4)$$

Таким образом, при делении комплексных чисел модули делятся, а аргументы вычитаются:

$$\left| \frac{z_1}{z_2} \right| = \frac{\rho_1}{\rho_2} = \frac{|z_1|}{|z_2|}, \quad (5)$$

$$\text{Arg} \frac{z_1}{z_2} = \text{Arg} z_1 - \text{Arg} z_2. \quad (6)$$

Для построения вектора, изображающего комплексное число $z_3 = \frac{z_1}{z_2}$, надо вектор $\vec{Oz}_1$ повернуть на угол φ_2 по часовой стрелке, если $\varphi_2 > 0$, и против, если $\varphi_2 < 0$, и «уменьшить» его длину в ρ_2 раз.

Деление комплексного числа z на i можно рассматривать как операцию поворота вектора $\vec{Oz}$ на угол $\frac{\pi}{2}$ по часовой стрелке.

Возведение в степень комплексных чисел. Возведение в степень n (где n — натуральное число) комплексного числа z рассматривается как n -кратное умножение z на самого себя.

Пусть $z = \rho \cos \varphi + i \sin \varphi$. Тогда

$$z^n = \underbrace{z \cdot z \cdot \dots \cdot z}_{n \text{ раз}} = \underbrace{\rho \cdot \rho \cdot \dots \cdot \rho}_{n \text{ раз}} [\underbrace{\cos (\varphi + \varphi + \dots + \varphi)}_{n \text{ раз}} + i \underbrace{\sin (\varphi + \varphi + \dots + \varphi)}_{n \text{ раз}}],$$

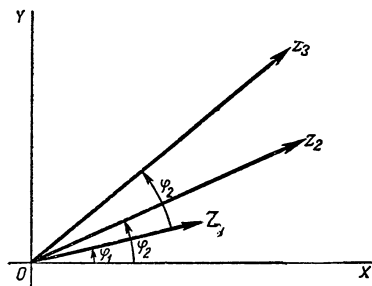


Рис. 241.

т. е.

$$z^n = \rho^n (\cos n\varphi + i \sin n\varphi). \quad (7)$$

Таким образом, при возведении комплексного числа в степень модуль возводится в эту степень, а аргумент умножается на число, равное показателю степени:

$$|z^n| = \rho^n = |z|^n, \quad (8)$$

$$\operatorname{Arg}(z^n) = n \operatorname{Arg} z. \quad (9)$$

Формулу (7) можно записать так:

$$[\rho (\cos \varphi + i \sin \varphi)]^n = \rho^n (\cos n\varphi + i \sin n\varphi). \quad (10)$$

В частном случае, когда $\rho = 1$, формула (10) имеет вид:

$$(\cos \varphi + i \sin \varphi)^n = \cos n\varphi + i \sin n\varphi. \quad (11)$$

Формула (11) называется *формулой Муавра*.

Пример. Выразить $\sin 3\varphi$ и $\cos 3\varphi$ через $\sin \varphi$ и $\cos \varphi$. Применим формулу Муавра при $n=3$:

$$(\cos \varphi + i \sin \varphi)^3 = \cos 3\varphi + i \sin 3\varphi,$$

$$\cos^3 \varphi + 3i \cos^2 \varphi \sin \varphi - 3 \cos \varphi \sin^2 \varphi - i \sin^3 \varphi = \cos 3\varphi + i \sin 3\varphi,$$

откуда

$$\cos 3\varphi = \cos^3 \varphi - 3 \cos \varphi \sin^2 \varphi,$$

$$\sin 3\varphi = 3 \cos^2 \varphi \sin \varphi - \sin^3 \varphi.$$

Извлечение корня из комплексных чисел. Пусть n — натуральное число. Корнем n -степени из комплексного числа $w = r(\cos \theta + i \sin \theta)$ называется комплексное число $z = \rho(\cos \varphi + i \sin \varphi)$, для которого $z^n = w$.

Обозначим его $z = \sqrt[n]{w}$.

По определению

$$\rho^n (\cos n\varphi + i \sin n\varphi) = r (\cos \theta + i \sin \theta). \quad (12)$$

Два комплексных числа в тригонометрической форме могут быть равны только тогда, когда равны их модули, а аргументы отличаются на целое кратное 2π :

$$\rho^n = r; \quad n\varphi = \theta + 2\pi k,$$

откуда

$$\rho = \sqrt[n]{r}, \quad (13)$$

$$\varphi = \frac{\theta + 2\pi k}{n}. \quad (14)$$

Следовательно,

$$\sqrt[n]{r (\cos \theta + i \sin \theta)} = \sqrt[n]{r} \left(\cos \frac{\theta + 2\pi k}{n} + i \sin \frac{\theta + 2\pi k}{n} \right). \quad (15)$$

Давая k значения $0, 1, 2, \dots, (n-1)$, получим n различных значений корня. Для значений $k=n, n+1, \dots$ или $k=-1, -2, \dots$ и т. д. значения корней будут повторять полученные ранее значения.

Например, при $k=0$ имеем $\sin \frac{\theta+0}{n} = \sin \frac{\theta}{n}$. При $k=n$ имеем $\sin \frac{\theta+2\pi n}{n} = \sin \left(\frac{\theta}{n} + 2\pi \right) = \sin \frac{\theta}{n}$ и т. д.

Пример. Найти все значения $\sqrt[5]{1}$. Так как $1 = 1(\cos 0 + i \sin 0)$, то $\sqrt[5]{1 (\cos 0 + i \sin 0)} = 1 \left(\cos \frac{0+2\pi k}{5} + i \sin \frac{0+2\pi k}{5} \right)$.

1) $k=0, z_0=1.$

2) $k=1, z_1 = \cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5}.$

$$3) k=2, z_2 = \cos \frac{4\pi}{5} + i \sin \frac{4\pi}{5}.$$

$$4) k=3, z_3 = \cos \frac{6\pi}{5} + i \sin \frac{6\pi}{5}.$$

$$5) k=4, z_4 = \cos \frac{8\pi}{5} + i \sin \frac{8\pi}{5}.$$

Дадим геометрическую интерпретацию решенного примера. Модули всех найденных значений $\sqrt[5]{1}$ равны единице. Следовательно, точки z_0, z_1, z_2, z_3, z_4 лежат на окружности, радиус которой $\rho=1$.

Углы $0, \frac{2\pi}{5}, \frac{4\pi}{5}, \frac{6\pi}{5}, \frac{8\pi}{5}$ определяют аргументы соответствующих точек.

Построим точки z_0, z_1, z_2, z_3, z_4 и проведем радиусы-векторы этих точек (рис. 242).

Если соединить точки z_0, z_1, z_2, z_3, z_4 прямыми, то получится правильный многоугольник с вершинами в точках $z_0, z_1, \dots, z_4$.

Обобщая данный частный случай, можно сказать, что геометрически точки, соответствующие различным значениям корня n -й степени из комплексного числа $w = r(\cos \theta + i \sin \theta)$, располагаются в вершинах правильного n -угольника с центром O , причем одна из вершин соответствует числу $\sqrt[n]{r} \left(\cos \frac{\theta}{n} + i \sin \frac{\theta}{n} \right)$.

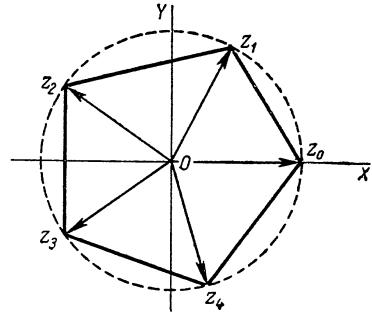


Рис. 242.

Раздел 3

ОБЩИЕ ПРИЕМЫ ИНТЕГРИРОВАНИЯ

§ 8. Интегрирование элементарных дробей

Элементарными дробями называются дроби следующих четырех типов:

I) $\frac{1}{ax+b}$, II) $\frac{1}{(ax+b)^m}$, III) $\frac{Mx+N}{ax^2+bx+c}$, IV) $\frac{Mx+N}{(ax^2+bx+c)^n}$, где m и n — натуральные числа ($m \geq 2, n \geq 2$) и $b^2 - 4ac < 0$.

Выясним, как интегрируются элементарные дроби.

Интегралы I) $\int \frac{dx}{ax+b}$ и II) $\int \frac{dx}{(ax+b)^m}$ подстановкой $t=ax+b$ сводятся к табличным интегралам:

$$I. \int \frac{dx}{ax+b} = \frac{1}{a} \int \frac{dt}{t} = \frac{1}{a} \ln |t| + C = \frac{1}{a} \ln |ax+b| + C. \quad (1)$$

$$II. \int \frac{dx}{(ax+b)^m} = \frac{1}{a} \int \frac{dt}{t^m} = -\frac{1}{a(m-1)t^{m-1}} + C = \\ = -\frac{1}{a(m-1)(ax+b)^{m-1}} + C. \quad (2)$$

III. Прежде чем рассмотреть вопрос о нахождении интеграла вида $\int \frac{Mx+N}{ax^2+bx+c} dx$, рассмотрим частный случай, когда $M=0, N=1$, т. е. рассмотрим интеграл

$$III, a. \int \frac{dx}{ax^2+bx+c}.$$

Преобразуем квадратный трехчлен ax^2+bx+c :

$$ax^2+bx+c = \frac{1}{4a} [4a^2x^2 + 4abx + b^2 + 4ac - b^2] = \\ = \frac{1}{4a} [(2ax+b)^2 + (4ac - b^2)].$$

Тогда данный интеграл можно переписать в следующей форме:

$$\int \frac{dx}{ax^2 + bx + c} = 4a \int \frac{dx}{(2ax + b)^2 + (4ac - b^2)}.$$

Обозначим $u = 2ax + b$ (тогда $du = 2adx$),

$$4ac - b^2 = s. \quad (3)$$

Заметим, что по условию $s > 0$. Тогда

$$\int \frac{dx}{ax^2 + bx + c} = \frac{4a}{2a} \int \frac{du}{u^2 + s} = \frac{2}{\sqrt{s}} \operatorname{arctg} \frac{u}{\sqrt{s}} + C = \frac{2}{\sqrt{s}} \operatorname{arctg} \frac{2ax + b}{\sqrt{s}} + C.$$

Таким образом, выделением квадрата двучлена интеграл III, а сводится к табличному интегралу.

Теперь рассмотрим интеграл III. Сначала преобразуем знаменатель так, как было сделано для частного случая, и применим подстановку (3). Тогда

$$\begin{aligned} \int \frac{Mx + N}{ax^2 + bx + c} dx &= 4a \int \frac{Mx + N}{(2ax + b)^2 + (4ac - b^2)} dx = \frac{4a}{2a} \int \frac{\frac{M(u-b)}{2a} + N}{u^2 + s} du = \\ &= \frac{1}{a} \int \frac{Mu - Mb + 2aN}{u^2 + s} du = \frac{M}{a} \int \frac{udu}{u^2 + s} + \frac{2aN - Mb}{a} \int \frac{du}{u^2 + s}. \end{aligned}$$

К первому из интегралов применим подстановку $t = u^2 + s$:

$$\int \frac{Mx + N}{ax^2 + bx + c} dx = \frac{M}{2a} \int \frac{dt}{t} + \frac{2aN - Mb}{a} \int \frac{du}{u^2 + s}.$$

Нахождение интеграла III сводится к двум табличным интегралам II, IX.

Пример 1.

$$\int \frac{7x - 2}{3x^2 - 5x + 4} dx.$$

Для решения примера 1 нам надо будет в знаменателе выделить квадрат двучлена. Для того чтобы упростить технические выкладки, полезно, если коэффициент при x^2 не является квадратом целого числа, умножить предварительно числитель и знаменатель на учетверенный коэффициент при x^2 (в данном случае на 12):

$$I = \int \frac{7x - 2}{3x^2 - 5x + 4} dx = \int \frac{84x - 24}{36x^2 - 60x + 48} dx = \int \frac{84x - 24}{(6x - 5)^2 + 23} dx.$$

Обозначим $u = 6x - 5$, тогда $du = 6dx$, $x = \frac{u+5}{6}$, и подставим в данный интеграл:

$$\begin{aligned} I &= \frac{1}{6} \int \frac{14u + 70 - 24}{u^2 + 23} du = \frac{7}{3} \int \frac{udu}{u^2 + 23} + \frac{23}{3} \int \frac{du}{u^2 + 23} = \frac{7}{6} \ln |u^2 + 23| + \\ &+ \frac{23}{3\sqrt{23}} \operatorname{arctg} \frac{u}{\sqrt{23}} + C = \frac{7}{6} \ln |36x^2 - 60x + 48| + \frac{\sqrt{23}}{3} \operatorname{arctg} \frac{6x - 5}{\sqrt{23}} + C = \\ &= \frac{7}{6} \ln |3x^2 - 5x + 4| + \frac{\sqrt{23}}{3} \operatorname{arctg} \frac{6x - 5}{\sqrt{23}} + C_1. \end{aligned}$$

Примечание 1. Если у квадратного трехчлена $ax^2 + bx + c$ $b^2 - 4ac > 0$, то дробь $\frac{Mx + N}{ax^2 + bx + c}$ по определению не является элементарной дробью, однако ее также можно интегрировать предложенным методом.

Пример 2. $I = \int \frac{5x - 3}{x^2 + 6x - 40} dx = \int \frac{5x - 3}{(x + 3)^2 - 49} dx$. Обозначим $u = x + 3$, тогда $du = dx$, $x = u - 3$,

$$\begin{aligned} I &= \int \frac{5u - 15 - 3}{u^2 - 49} du = 5 \int \frac{udu}{u^2 - 49} - 18 \int \frac{du}{u^2 - 49} = \\ &= \frac{5}{2} \ln |u^2 - 49| - \frac{18}{14} \ln \left| \frac{u - 7}{u + 7} \right| + C = \frac{5}{2} \ln |x^2 + 6x - 40| - \frac{9}{7} \ln \left| \frac{x - 4}{x + 10} \right| + C. \end{aligned}$$

Примечание 2. Интегралы от иррациональных функций вида $\frac{Mx+N}{\sqrt{ax^2+bx+c}}$ также можно брать методом, аналогичным указанному для интеграла III.

Пример 3.

$$I = \int \frac{3x+4}{\sqrt{7-x^2+6x}} dx = \int \frac{3x+4}{\sqrt{16-(x-3)^2}} dx.$$

Обозначим $u = x - 3$, тогда $du = dx$, $x = u + 3$, $I = \int \frac{3u+9+4}{\sqrt{16-u^2}} du = 3 \int \frac{udu}{\sqrt{16-u^2}} +$
 $+ 13 \int \frac{du}{\sqrt{16-u^2}} = -3\sqrt{16-u^2} + 13 \arcsin \frac{u}{4} + C = -3\sqrt{7-x^2+6x} +$
 $+ 13 \arcsin \frac{x-3}{4} + C.$

IV. Прежде чем рассмотреть интеграл вида $\int \frac{Mx+N}{(ax^2+bx+c)^n} dx$, рассмотрим случай, когда $M=0$, $N=1$, т. е. интеграл вида $\int \frac{dx}{(ax^2+bx+c)^n}$.

Как было показано, выделением квадрата двучлена и соответствующей заменой квадратный трехчлен ax^2+bx+c может быть выражен через квадратный двучлен u^2+s .

Поэтому нахождение интеграла $\int \frac{dx}{(ax^2+bx+c)^n}$ сводится к нахождению интеграла

$$\int \frac{du}{(u^2+s)^n}. \quad (4)$$

Преобразуем предварительно интеграл (4)

$$\int \frac{du}{(u^2+s)^n} = \frac{1}{s} \int \frac{s+u^2-u^2}{(u^2+s)^n} du = \frac{1}{s} \int \frac{du}{(u^2+s)^{n-1}} - \frac{1}{s} \int \frac{u^2 du}{(u^2+s)^n}. \quad (5)$$

Интеграл $\int \frac{u^2 du}{(u^2+s)^n}$ будем брать по частям; для этого обозначим

$$dv_1 = \frac{udu}{(u^2+s)^n}, \quad u_1 = u, \quad \text{тогда } du_1 = du; \quad v_1 = \int \frac{udu}{(u^2+s)^n} =$$

 $= -\frac{1}{2(n-1)(u^2+s)^{n-1}},$
 $\int \frac{u^2 du}{(u^2+s)^n} = -\frac{u}{(2n-2)(u^2+s)^{n-1}} + \frac{1}{2n-2} \int \frac{du}{(u^2+s)^{n-1}}. \quad (6)$

Подставим найденное значение интеграла в равенство (5):

$$\int \frac{du}{(u^2+s)^n} = \frac{1}{s} \int \frac{du}{(u^2+s)^{n-1}} + \frac{u}{s(2n-2)(u^2+s)^{n-1}} - \frac{1}{s(2n-2)} \int \frac{du}{(u^2+s)^{n-1}}.$$

Приведя подобные интегралы, получим:

$$\int \frac{du}{(u^2+s)^n} = \frac{u}{s(2n-2)(u^2+s)^{n-1}} + \frac{2n-3}{s(2n-2)} \int \frac{du}{(u^2+s)^{n-1}}. \quad (7)$$

Получившееся равенство позволяет интеграл $\int \frac{du}{(u^2+s)^n}$ выразить через интеграл $\int \frac{du}{(u^2+s)^{n-1}}$, т. е. понизить степень знаменателя в подынтегральном выражении на единицу. Применяя формулу (7) последовательно $(n-1)$ раз, придем к табличному интегралу $\int \frac{du^{\frac{1}{2}}}{u^2+s}$.

¹⁾ Формулу (7) можно применять и тогда, когда $s = 4ac - b^2 < 0$. В этом случае указанные преобразования в конечном итоге приведут к табличному интегралу X:

$$\int \frac{dx}{x^2 - a^2}.$$

Формула (7) называется *рекуррентной*.

Пример 4. $\int \frac{dx}{(x^2+5)^3}$. Применим формулу (7). В данном случае $u=x$, $s=5$, $n=3$.

$$\begin{aligned} \int \frac{dx}{(x^2+5)^3} &= \frac{x}{5 \cdot 4 (x^2+5)^2} + \frac{3}{5 \cdot 4} \int \frac{dx}{(x^2+5)^2} = \frac{x}{20 (x^2+5)^2} + \\ &+ \frac{3}{20} \left\{ \frac{x}{5 \cdot 2 (x^2+5)} + \frac{1}{5 \cdot 2} \int \frac{dx}{x^2+5} \right\} = \frac{x}{20 (x^2+5)^2} + \frac{3x}{200 (x^2+5)} + \\ &+ \frac{3}{200 \sqrt{5}} \operatorname{arctg} \frac{x}{\sqrt{5}} + C. \end{aligned}$$

В общем случае интеграл $\int \frac{Mx+N}{(ax^2+bx+c)^n} dx$ берется следующим образом: преобразуют квадратный трехчлен и делают подстановку $u=2ax+b$ ($du=2adx$, $x=\frac{u-b}{2a}$), $s=4ac-b^2$. Тогда

$$\begin{aligned} \int \frac{Mx+N}{(ax^2+bx+c)^n} dx &= (4a)^n \int \frac{Mx+N}{[(2ax+b)^2+(4ac-b^2)]^n} dx = \\ &= \frac{(4a)^n}{2a} \int \frac{\frac{M(u-b)}{2a} + N}{(u^2+s)^n} du = \frac{(4a)^n}{2a} \left\{ \frac{M}{2a} \int \frac{udu}{(u^2+s)^n} + \right. \\ &\quad \left. + \frac{2aN-Mb}{2a} \int \frac{du}{(u^2+s)^n} \right\}. \end{aligned}$$

В последнем равенстве первый из интегралов подстановкой $t=u^2+s$ приводится к табличному интегралу $\int \frac{dt}{t^n}$, а второй берется по рекуррентной формуле (7).

Пример 5.

$$\begin{aligned} I &= \int \frac{3x+5}{(x^2-4x+7)^2} dx = \int \frac{3x+5}{[(x-2)^2+3]^2} dx; \\ &\quad u=x-2, du=dx, x=u+2; \\ I &= \int \frac{3u+6+5}{(u^2+3)^2} du = 3 \int \frac{udu}{(u^2+3)^2} + 11 \int \frac{du}{(u^2+3)^2}; \\ &\quad t=u^2+3, dt=2udu; \\ I &= \frac{3}{2} \int \frac{dt}{t^2} + 11 \left\{ \frac{u}{3 \cdot 2 (u^2+3)} + \frac{1}{3 \cdot 2} \int \frac{du}{u^2+3} \right\} = -\frac{3}{2t} + \frac{11u}{6(u^2+3)} + \\ &\quad + \frac{11}{6\sqrt{3}} \operatorname{arctg} \frac{u}{\sqrt{3}} + C = -\frac{3}{2(x^2-4x+7)} + \frac{11(x-2)}{6(x^2-4x+7)} + \\ &\quad + \frac{11}{6\sqrt{3}} \operatorname{arctg} \frac{x-2}{\sqrt{3}} + C. \end{aligned}$$

§ 9. Интегрирование рациональных функций

Разложение многочлена на линейные множители. Рассмотрим многочлен $P(x)$ степени n . Каждое действительное или комплексное число x , которое обращает $P(x)$ в нуль, называется *корнем* этого многочлена. Следовательно, корни многочлена $P(x)$ являются решениями уравнения

$$P(x)=0. \quad (1)$$

В высшей алгебре доказывается следующая основная теорема — теорема Гаусса, которую примем без доказательства: «Всякий многочлен степени не ниже первой имеет по крайней мере один действительный или мнимый корень».

В элементарной алгебре была доказана теорема Безу о том, что «если x_1 является корнем многочлена, то этот многочлен делится без остатка на $x-x_1$ ».

Пусть $P_n(x)$ — многочлен степени n :

$$P_n(x) = C_n x^n + C_{n-1} x^{n-1} + \dots + C_1 x + C_0. \quad (2)$$

На основании теоремы Гаусса можно утверждать, что многочлен $P_n(x)$ имеет по крайней мере один корень x_1 . Тогда, применяя теорему Безу, можно записать:

$$P_n(x) = (x - x_1) P_{n-1}(x), \quad (3)$$

где $P_{n-1}(x)$ — многочлен степени $n - 1$. Продолжая аналогичные рассуждения, применяя последовательно теоремы Гаусса и Безу, можно получить, что

$$P_n(x) = (x - x_1)(x - x_2) \dots (x - x_n) P_0(x), \quad (4)$$

где $P_0(x)$ — многочлен нулевой степени. Легко видеть, что многочлен $P_0(x)$ равен C_n .

Таким образом приходим к выводу, что всякий многочлен $P_n(x)$ степени n может быть представлен в виде произведения линейных множителей вида $x - x_k$ и постоянного множителя C_n — коэффициента при старшей степени x многочлена $P_n(x)$:

$$P_n(x) = C_n (x - x_1)(x - x_2) \dots (x - x_n). \quad (5)$$

Если среди корней многочлена $x_1, x_2, \dots, x_n$ число α встречается m раз, то это число α называется *m-кратным корнем* многочлена. При $m = 1$ число α называется *простым корнем*. Собирая вместе множители, соответствующие одинаковым корням, получим, что всякий многочлен может быть записан в форме:

$$P_n(x) = C_n (x - x_1)^k (x - x_2)^l \dots (x - x_q)^n, \quad (6)$$

причем любой из показателей $k, l, \dots, m$ может в частном случае равняться единице, а

$$k + l + \dots + m = n. \quad (7)$$

Разложение многочлена с действительными коэффициентами на линейные и квадратичные множители. Предварительно докажем следующую теорему: если комплексное число $z = u + iv$ является корнем многочлена $P_n(x)$ с действительными коэффициентами $C_n, C_{n-1}, \dots, C_0$, то и сопряженное ему число $\bar{z} = u - iv$ также будет корнем этого многочлена. Действительно, так как $z = u + iv$ является корнем многочлена $P_n(x)$, то

$$P_n(z) = C_n z^n + C_{n-1} z^{n-1} + \dots + C_0 = 0.$$

На основании свойств комплексных чисел можно утверждать, что

$$\begin{aligned} P_n(\bar{z}) &= C_n \bar{z}^n + C_{n-1} \bar{z}^{n-1} + \dots + C_0 = C_n \overline{z^n} + C_{n-1} \overline{z^{n-1}} + \dots + C_0 = \\ &= \overline{C_n z^n + C_{n-1} z^{n-1} + \dots + C_0} = \overline{P_n(z)} = 0, \end{aligned} \quad (8)$$

т. е. $\bar{z}$ — корень многочлена $P(z)$. Можно показать, что корни z и $\bar{z}$ имеют одинаковую кратность.

Рассмотрим произведение множителей с попарно сопряженными мнимыми корнями.

Пусть $z = u + iv$ и $\bar{z} = u - iv$ являются корнями некоторого многочлена, тогда произведение $(x - z)(x - \bar{z})$ можно преобразовать следующим образом:

$$\begin{aligned} (x - z)(x - \bar{z}) &= [x - (u + iv)][x - (u - iv)] = \\ &= [(x - u) - iv][(x - u) + iv] = (x - u)^2 + v^2 = x^2 - 2xu + u^2 + v^2. \end{aligned} \quad (9)$$

Обозначим

$$-2u = p; \quad u^2 + v^2 = q. \quad (10)$$

Тогда

$$(x - z)(x - \bar{z}) = x^2 + px + q, \quad (11)$$

где p и q — действительные числа $\frac{p^2}{4} - q < 0$. Следовательно, произведение линейных множителей, соответствующих сопряженным корням, можно заменить квадратным трехчленом с действительными коэффициентами.

На основании вышеизложенного всякий многочлен с действительными коэффициентами может быть представлен как произведение линейных и квадратичных множителей с действительными коэффициентами следующим образом:

$$P_n(x) = C_n (x-a)^\alpha (x-b)^\beta \dots (x^2+px+q)^\lambda \dots (x^2+rx+s)^\mu, \quad (12)$$

где $\alpha + \beta + \dots + 2\lambda + \dots + 2\mu = n$, $\frac{p^2}{4} - q < 0$, $\dots$, $\frac{r^2}{4} - s < 0$.

Разложение рациональной дроби на элементарные. Ранее указывалось, что рациональную функцию можно рассматривать как отношение двух многочленов¹⁾. Рациональная дробь называется *правильной*, если степень многочлена в числителе меньше степени многочлена в знаменателе. В противном случае рациональная дробь называется *неправильной*. Выделяя из неправильной дроби ее целую часть (путем деления числителя на знаменатель), всегда можно представить ее в виде суммы многочлена и правильной рациональной дроби.

Пример. $\frac{x^2}{x^2+3} = 1 + \frac{-3}{x^2+3}$.

Так как интегрирование многочлена не представляет трудностей, то достаточно научиться интегрировать правильные рациональные дроби. Сформулированная ниже теорема позволяет свести интегрирование любой правильной рациональной дроби к интегрированию элементарных дробей.

Теорема. Если $R(x) = \frac{Q(x)}{P(x)}$ — правильная рациональная дробь, знаменатель $P(x)$ которой представлен в виде произведения линейных и квадратичных множителей (с действительными коэффициентами)

$$P(x) = (x-a)^\alpha \dots (x-b)^\beta (x^2+px+q)^\lambda \dots (x^2+rx+s)^\mu, \quad (13)$$

то эта дробь может быть разложена на элементарные дроби по следующей схеме:

$$\begin{aligned} \frac{Q(x)}{P(x)} = & \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \dots + \frac{A_\alpha}{(x-a)^\alpha} + \dots + \frac{B_1}{x-b} + \frac{B_2}{(x-b)^2} + \\ & + \dots + \frac{B_\beta}{(x-b)^\beta} + \frac{M_1x+N_1}{x^2+px+q} + \frac{M_2x+N_2}{(x^2+px+q)^2} + \dots + \frac{M_\lambda x+N_\lambda}{(x^2+px+q)^\lambda} + \\ & + \dots + \frac{R_1x+S_1}{x^2+rx+s} + \frac{R_2x+S_2}{(x^2+rx+s)^2} + \dots + \frac{R_\mu x+S_\mu}{(x^2+rx+s)^\mu}, \end{aligned} \quad (14)$$

где $A_1, \dots, A_\alpha, B_1, \dots, B_\beta, \dots, R_1, S_1, \dots, R_\mu, S_\mu$ — некоторые действительные числа³⁾.

Практически разложение конкретной правильной рациональной дроби $\frac{Q(x)}{P(x)}$ на сумму элементарных дробей обычно производят *методом неопределенных коэффициентов*. Для этого разлагают знаменатель $P(x)$ на произведение линейных и квадратичных множителей, записывают разложение дроби $\frac{Q(x)}{P(x)}$ по схеме (14) с неопределенными коэффициентами, затем при-

¹⁾ Не ограничивая общности, можно считать, что эти многочлены не имеют общих корней.

²⁾ Постоянный множитель C_n — коэффициент при старшей степени x многочлена $P(x)$ — всегда можно вынести за знак интеграла, поэтому в разложении (13) полагаем C_n равным единице.

³⁾ Доказательство этой теоремы читатель найдет в более полных курсах математического анализа.

водят элементарные дроби к общему знаменателю $P(x)$ и приравнивают многочлен, получившийся в числителе, многочлену $Q(x)$.

Для того чтобы два многочлена были тождественно равны, необходимо и достаточно, чтобы коэффициенты при одинаковых степенях x у них были равны. Учитывая это замечание, приравниваем коэффициенты при одинаковых степенях в левой и правой частях равенства, получая тем самым систему алгебраических уравнений для нахождения неопределенных коэффициентов. Существование решения такой системы вытекает из сформулированной выше теоремы.

Пример. Дробь $\frac{9x^3 - 30x^2 + 28x - 88}{(x^2 - 6x + 8)(x^2 + 4)}$ — правильная.

Так как $(x^2 - 6x + 8)(x^2 + 4) = (x - 2)(x - 4)(x^2 + 4)$, то

$$\frac{9x^3 - 30x^2 + 28x - 88}{(x^2 - 6x + 8)(x^2 + 4)} = \frac{A}{x - 2} + \frac{B}{x - 4} + \frac{Cx + D}{x^2 + 4},$$

где A , B , C и D — пока неопределенные коэффициенты.

Приведя дроби в правой части равенства к общему знаменателю и приравнявая многочлен, получившийся в числителе, многочлену $9x^3 - 30x^2 + 28x - 88$, получим:

$$A(x - 4)(x^2 + 4) + B(x - 2)(x^2 + 4) + (Cx + D)(x^2 - 6x + 8) = 9x^3 - 30x^2 + 28x - 88, \quad (15)$$

т. е.

$$(A + B + C)x^3 + (-4A - 2B - 6C + D)x^2 + (4A + 4B + 8C - 6D)x + (-16A - 8B + 8D) = 9x^3 - 30x^2 + 28x - 88.$$

Приравниваем коэффициенты при одинаковых степенях x .

$$\begin{array}{l|l} x^3 & A + B + C = 9, \\ x^2 & -4A - 2B - 6C + D = -30, \\ x & 4A + 4B + 8C - 6D = 28, \\ x^0 & -16A - 8B + 8D = -88. \end{array}$$

Решая эту систему уравнений, находим:

$$A = 5, \quad B = 3, \quad C = 1, \quad D = 2,$$

а значит,

$$\frac{9x^3 - 30x^2 + 28x - 88}{(x^2 - 6x + 8)(x^2 + 4)} = \frac{5}{x - 2} + \frac{3}{x - 4} + \frac{x + 2}{x^2 + 4}.$$

Вместо того, чтобы сравнивать коэффициенты при одинаковых степенях x , можно при нахождении неопределенных коэффициентов дать переменной x ряд конкретных произвольных значений (столько, сколько неопределенных коэффициентов) и получить таким образом систему уравнений относительно неопределенных коэффициентов. Этот прием называется *методом произвольных значений*.

Так, в рассмотренном выше примере, подставляя в равенство (15) последовательно $x = 2$, $x = 4$, $x = 0$, $x = 1$, можно получить систему уравнений:

$$\begin{aligned} -16A &= -80, \\ 40B &= 120, \\ -16A - 8B + 8D &= -88, \\ -15A - 5B + 3C + 3D &= -81, \end{aligned}$$

откуда $A = 5$, $B = 3$, $C = 1$, $D = 2$.

Иногда для нахождения из равенства (15) неопределенных коэффициентов удобно применять комбинацию указанных выше методов, давая x ряд произвольных значений и приравнявая коэффициенты при некоторых степенях x .

В рассмотренном примере удобно, например, положив $x=2$ и $x=4$, найти $A=5$, $B=3$, а затем, приравняв коэффициенты многочленов при x^3 и x^2 , получить уравнения:

$$A + B + C = 9 \text{ и } -4A - 2B - 6C + D = -30,$$

откуда определить $C = 9 - 5 - 3 = 1$, $D = -30 + 4A + 2B + 6C = 2$.

Обобщая результаты предшествующих разделов, заключаем, что всякая рациональная дробь может быть проинтегрирована в элементарных функциях.

Сформулируем теперь общее правило для интегрирования рациональной функции. Для того чтобы проинтегрировать рациональную дробь (рациональную функцию), надо:

1. Если рассматриваемая дробь $\frac{Q(x)}{P(x)}$ неправильная, то выделить из нее целую часть: $\frac{Q(x)}{P(x)} = T(x) + \frac{Q_1(x)}{P(x)}$, где $T(x)$ — многочлен, $\frac{Q_1(x)}{P(x)}$ — правильная рациональная дробь.

2. Правильную рациональную дробь $\frac{Q_1(x)}{P(x)}$ представить в виде суммы элементарных дробей по схеме (14).

3. Представить интеграл от данной дроби в виде суммы интегралов от целой части и от элементарных дробей и взять эти интегралы.

Пример 1. $\int \frac{6x^5 - 8x^4 - 25x^3 + 20x^2 - 76x - 7}{3x^3 - 4x^2 - 17x + 6} dx = I.$

Подинтегральная дробь неправильная. Поэтому выделим сначала целую часть и проинтегрируем ее, а в оставшейся правильной дроби разложим знаменатель на линейные множители:

$$\begin{aligned} I &= \int \left(2x^2 + 3 + \frac{20x^2 - 25x - 25}{3x^3 - 4x^2 - 17x + 6} \right) dx = \\ &= \frac{2}{3} x^3 + 3x + \int \frac{20x^2 - 25x - 25}{(x+2)(x-3)(3x-1)} dx. \end{aligned}$$

Выписываем оставшуюся подинтегральную правильную дробь и представляем ее в виде суммы элементарных дробей с неопределенными коэффициентами:

$$\frac{20x^2 - 25x - 25}{(x+2)(x-3)(3x-1)} = \frac{A}{x+2} + \frac{B}{x-3} + \frac{C}{3x-1}. \quad (16)$$

Для нахождения коэффициентов A , B , C в равенстве (16) приведем дроби к общему знаменателю и приравняем многочлены, полученные в числителях:

$$20x^2 - 25x - 25 = A(x-3)(3x-1) + B(x+2)(3x-1) + C(x+2)(x-3). \quad (17)$$

Для нахождения коэффициентов A , B и C будем давать x произвольные значения. Этот метод быстро приводит к цели, если в качестве произвольных значений брать корни многочлена, стоящего в знаменателе, т. е. в данном случае -2 , 3 , $\frac{1}{3}$.

Подставляя последовательно эти значения в равенство (17), получим:

$$\begin{aligned} x &= -2, & 105 &= 35A, & A &= 3, \\ x &= 3, & 80 &= 40B, & B &= 2, \\ x &= \frac{1}{3}, & -280 &= -56C, & C &= 5. \end{aligned}$$

Теперь, найдя коэффициенты A , B , C , на основании равенства (16) представим интеграл от данной дроби в виде суммы интегралов от элементарных дробей и проинтегрируем их:

$$\begin{aligned} \int \frac{20x^2 - 25x - 25}{(x+2)(x-3)(3x-1)} dx &= 3 \int \frac{dx}{x+2} + 2 \int \frac{dx}{x-3} + 5 \int \frac{dx}{3x-1} = \\ &= 3 \ln |x+2| + 2 \ln |x-3| + \frac{5}{3} \ln |3x-1| + C. \end{aligned}$$

Окончательно имеем:

$$I = \frac{2}{3} x^3 + 3x + 3 \ln |x+2| + 2 \ln |x-3| + \frac{5}{3} \ln |3x-1| + C.$$

Пример 2. $\int \frac{8x^3 - 12x^2 + 2x + 10}{(x+3)(x-1)^3} dx.$

Подинтегральная дробь правильная и знаменатель разложен на простейшие множители (линейные, простые и кратные). Будем решать пример 2 так же, как пример 1:

$$\begin{aligned} \frac{8x^3 - 12x^2 + 2x + 10}{(x+3)(x-1)^3} &= \frac{A}{x+3} + \frac{B}{x-1} + \frac{C}{(x-1)^2} + \frac{D}{(x-1)^3}, \\ 8x^3 - 12x^2 + 2x + 10 &= A(x-1)^3 + B(x+3)(x-1)^2 + C(x+3)(x-1) + \\ &\quad + D(x+3), \\ x=1, \quad 8 &= 4D, \quad D=2, \\ x=-3, \quad -320 &= -64A, \quad A=5. \end{aligned}$$

По исчерпанию всех значений, являющихся корнями многочлена, стоящего в знаменателе, следует брать значения x , наиболее простые для подсчета. Возьмем, например, значения x , равные 0 и 2:

$$\begin{aligned} x=0, \quad 10 &= -5 + 3B - 3C + 6, \quad B - C = 3, \\ x=2, \quad 30 &= 5 + 5B + 5C + 10, \quad B + C = 3, \end{aligned}$$

откуда $B=3, C=0$.

$$\begin{aligned} \int \frac{8x^3 - 12x^2 + 2x + 10}{(x+3)(x-1)^3} dx &= 5 \int \frac{dx}{x+3} + 2 \int \frac{dx}{(x-1)^3} + 3 \int \frac{dx}{x-1} = \\ &= 5 \ln |x+3| - \frac{1}{(x-1)^2} + 3 \ln |x-1| + C. \end{aligned}$$

Пример 3. $\int \frac{8x^3 - 2x^2 + 13x - 2}{(x^2+2)(x^2+1)} dx.$

Подинтегральная дробь правильная. Знаменатель разложен на простейшие множители, в данном случае простые, квадратичные.

Разложим подинтегральную дробь на элементарные:

$$\frac{8x^3 - 2x^2 + 13x - 2}{(x^2+2)(x^2+1)} = \frac{Ax+B}{x^2+2} + \frac{Cx+D}{x^2+1}. \quad (18)$$

В этом примере для нахождения коэффициентов A, B, C, D применим метод сравнения коэффициентов. Для этого в равенстве (18) освободимся от знаменателя и в правой части объединим члены при x в одинаковых степенях.

$$\begin{aligned} 8x^3 - 2x^2 + 13x - 2 &= Ax^3 + Bx^2 + Ax + B + \\ &\quad + Cx^3 + Dx^2 + 2Cx + 2D, \\ 8x^3 - 2x^2 + 13x - 2 &= (A+C)x^3 + (B+D)x^2 + \\ &\quad + (A+2C)x + B+2D. \end{aligned} \quad (19)$$

В тождестве (19) коэффициенты при одинаковых степенях x будут равны, следовательно, получим систему уравнений:

$$A+C=8, \quad B+D=-2, \quad A+2C=13, \quad B+2D=-2,$$

из которой определим коэффициенты: $A=3, B=-2, C=5, D=0$.

$$\begin{aligned} \int \frac{8x^3 - 2x^2 + 13x - 2}{(x^2+2)(x^2+1)} dx &= \int \frac{3x-2}{x^2+2} dx + 5 \int \frac{xdx}{x^2+1} = \\ &= \frac{3}{2} \ln |x^2+2| - \frac{2}{\sqrt{2}} \operatorname{arctg} \frac{x}{\sqrt{2}} + \frac{5}{2} \ln |x^2+1| + C. \end{aligned}$$

Пример 4.

$$\int \frac{3x^4 + 14x^2 + 7x + 15}{(x+3)(x^2+2)^2} dx.$$

Подинтегральная дробь правильная. Знаменатель разложен на простейшие множители, из которых один простой линейный, другой — кратный квадратичный. Раскладываем подинтегральную дробь на элементарные.

$$\frac{3x^4 + 14x^2 + 7x + 15}{(x+3)(x^2+2)^2} = \frac{A}{x+3} + \frac{Bx+D}{x^2+2} + \frac{Mx+N}{(x^2+2)^2}.$$

Для нахождения коэффициентов сначала дадим x ряд значений (до исчерпания всех значений действительных корней знаменателя), а затем перейдем к сравнению коэффициентов:

$$3x^4 + 14x^2 + 7x + 15 = A(x^2+2)^2 + (Bx+D)(x+3)(x^2+2) + (Mx+N)(x+3).$$

При $x = -3$ имеем $363 = 121A$, $A = 3$.
Далее найдем, что

$$2x^2 + 7x + 3 = Bx^4 + 3Bx^3 + 2Bx + 6Bx + Dx^3 + 3Dx^2 + 2Dx + \\ + 6D + Mx^2 + Nx + 3Mx + 3N, \\ \text{откуда } B = 0, D = 0, M = 2, N = 1.$$

$$\int \frac{3x^4 + 14x^2 + 7x + 15}{(x+3)(x^2+2)^2} dx = 3 \int \frac{dx}{x+3} + 2 \int \frac{x dx}{(x^2+2)^2} + \int \frac{dx}{(x^2+2)^2} = \\ = 3 \ln(x+3) - \frac{1}{x^2+2} + \frac{x}{4(x^2+2)} + \frac{1}{4\sqrt{2}} \operatorname{arctg} \frac{x}{\sqrt{2}} + C.$$

Точно указать, какой из приемов нахождения коэффициентов рациональнее применять, нельзя, однако можно рекомендовать в случае линейных простых множителей пользоваться методом произвольных значений; в случае квадратичных простых множителей — методом сравнения коэффициентов; в более сложных примерах — комбинацией этих методов.

§ 10. Интегрирование тригонометрических выражений

Введем некоторые определения и условные обозначения. Будем называть *многочленом* или *целой рациональной функцией относительно u и v* сумму конечного числа слагаемых вида $Cu^m v^k$, где m и k — натуральные, а C — действительные числа.

Рациональной функцией относительно u и v называется отношение двух многочленов относительно u и v .

Например, выражение $\frac{\sin^3 x + \cos x}{2 \sin x \cos^2 x + 3}$ является рациональной функцией относительно $\sin x$ и $\cos x$.

В дальнейшем символом $R(u, v)$ будем обозначать рациональную функцию относительно переменных u и v . Будем говорить, что функция $R(u, v)$ — *нечетная относительно u* , если при замене u на $-u$ функция $R(u, v)$ меняет знак, т. е. $R(-u, v) = -R(u, v)$.

Аналогично, если $R(u, -v) = -R(u, v)$, то будем говорить, что такая функция *нечетна относительно v* . Будем говорить, что функция $R(u, v)$ *четна относительно u и v* , если при одновременной замене u на $-u$, v на $-v$ функция не изменяется, т. е.

$$R(-u, -v) = R(u, v).$$

Переходя к рассмотрению приемов интегрирования тригонометрических выражений, предварительно заметим, что любая тригонометрическая функция аргумента x рационально выражается через тангенс половинного угла. Действительно,

$$\sin x = \frac{2 \sin \frac{x}{2} \cos \frac{x}{2}}{\cos^2 \frac{x}{2} + \sin^2 \frac{x}{2}} = \frac{2 \operatorname{tg} \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}}, \quad (1)$$

$$\cos x = \frac{\cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}}{\cos^2 \frac{x}{2} + \sin^2 \frac{x}{2}} = \frac{1 - \operatorname{tg}^2 \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}}, \quad (2)$$

а так как $\operatorname{tg} x$, $\operatorname{ctg} x$, $\sec x$, $\operatorname{cosec} x$ рационально выражаются через $\sin x$ и $\cos x$, то они рационально будут выражаться и через $\operatorname{tg} \frac{x}{2}$.

Теорема 1. Интеграл вида $\int R(\sin x, \cos x) dx$ подстановкой $t = \operatorname{tg} \frac{x}{2}$ приводится к интегралу от рациональной функции t , который, как было указано выше, всегда выражается в элементарных функциях.

Действительно, пусть

$$t = \operatorname{tg} \frac{x}{2}. \quad (3)$$

Тогда на основании равенств (1) и (2)

$$\sin x = \frac{2t}{1+t^2}, \quad (4)$$

$$\cos x = \frac{1-t^2}{1+t^2}. \quad (5)$$

Определяя x и dx из равенства (3), найдем, что

$$x = 2 \operatorname{arctg} t; \quad dx = \frac{2dt}{1+t^2}. \quad (6)$$

В подинтегральное выражение вместо $\sin x$, $\cos x$ и dx подставим их значения, определяемые равенствами (4), (5) и (6):

$$\int R(\sin x, \cos x) dx = \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right) \frac{2}{1+t^2} dt = \int r(t) dt,$$

где $r(t)$ — рациональная функция t . Так как подстановкой $t = \operatorname{tg} \frac{x}{2}$ интеграл от любой рациональной относительно $\sin x$ и $\cos x$ тригонометрической функции приводится к интегралу от рациональной функции, то подстановка (3) называется *универсальной подстановкой*.

Пример.
$$\int \frac{dx}{9 + 8 \cos x + \sin x}.$$

Применим универсальную подстановку $t = \operatorname{tg} \frac{x}{2}$:

$$\begin{aligned} \int \frac{dx}{9 + 8 \cos x + \sin x} &= \int \frac{2dt}{(1+t^2) \left[9 + \frac{8(1-t^2)}{1+t^2} + \frac{2t}{1+t^2} \right]} = \\ &= 2 \int \frac{dt}{t^2 + 2t + 17} = 2 \int \frac{dt}{(t+1)^2 + 16} = \frac{1}{2} \operatorname{arctg} \frac{t+1}{4} + C = \\ &= \frac{1}{2} \operatorname{arctg} \frac{\operatorname{tg} \frac{x}{2} + 1}{4} + C. \end{aligned}$$

Следует заметить, что хотя универсальная подстановка для интегралов вида $\int R(\sin x, \cos x) dx$ всегда приводит к цели, однако ее используют сравнительно редко, так как она часто приводит к громоздким выкладкам. Рассмотрим другие приемы интегрирования тригонометрических выражений, которые в частных случаях быстрее приводят к цели.

Теорема 2. Если функция $R(\sin x, \cos x)$ нечетна относительно $\cos x$, то интеграл $\int R(\sin x, \cos x) dx$ подстановкой $t = \sin x$ приводится к интегралу от рациональной функции.

Указанный в теореме 2 интеграл преобразуем:

$$\int R(\sin x, \cos x) dx = \int \frac{R(\sin x, \cos x)}{\cos x} \cos x dx.$$

Так как функция $R(\sin x, \cos x)$ нечетна относительно $\cos x$, то выражение $\frac{R(\sin x, \cos x)}{\cos x}$ четно относительно $\cos x$ и может содержать $\cos x$ лишь в четных степенях. Следовательно, это выражение является рациональной функцией относительно $\sin x$:

$$\frac{R(\sin x, \cos x)}{\cos x} = r(\sin x),$$

т. е.

$$\int \frac{R(\sin x, \cos x)}{\cos x} \cos x dx = \int r(\sin x) \cos x dx.$$

К последнему интегралу применим подстановку $t = \sin x$, тогда $dt = \cos x dx$:

$$\int R(\sin x, \cos x) dx = \int r(\sin x) \cos x dx = \int r(t) dt.$$

Таким образом, указанная подстановка привела данный интеграл к интегралу от рациональной функции относительно t .

Теорема 3. Если $R(\sin x, \cos x)$ имеет нечетный характер относительно $\sin x$, то подстановкой $t = \cos x$ интеграл $\int R(\sin x, \cos x) dx$ приводится к интегралу от рациональной функции.

Теорема 3 доказывается так же, как теорема 2.

Пример 1. $\int \frac{\cos^3 x}{2 + \sin x} dx$. Подинтегральная функция нечетна относительно $\cos x$. Применяя подстановку $t = \sin x$ (тогда $dt = \cos x dx$), получим:

$$\begin{aligned} \int \frac{\cos^3 x}{2 + \sin x} dx &= \int \frac{(1 - \sin^2 x)}{2 + \sin x} \cos x dx = \int \frac{1 - t^2}{2 + t} dt = \\ &= \int \left(-t + 2 - \frac{3}{2+t} \right) dt = -\frac{t^2}{2} + 2t - 3 \ln |2 + t| + C = -\frac{1}{2} \sin^2 x + \\ &\quad + 2 \sin x - 3 \ln |2 + \sin x| + C. \end{aligned}$$

Примечание 1. Теоремы 2 и 3 обычно применяются при интегрировании произведения $\sin^m x \cos^n x$ натуральных степеней $\cos x$ и $\sin x$, когда или m или n (или m и n) — нечетные.

Пример 2. $\int \cos^3 x \sin^5 x dx$.

Подинтегральная функция содержит $\sin x$ в нечетной степени, поэтому применяем подстановку $t = \cos x$; тогда $dt = -\sin x dx$,

$$\begin{aligned} \int \cos^3 x \sin^5 x dx &= \int \cos^2 x (1 - \cos^2 x)^2 \sin x dx = -\int t^2 (1 - t^2)^2 dt = \\ &= -\int t^2 (1 - 2t^2 + t^4) dt = -\frac{1}{3} t^3 + \frac{2}{5} t^5 - \frac{1}{7} t^7 + C = -\frac{1}{3} \cos^3 x + \\ &\quad + \frac{2}{5} \cos^5 x - \frac{1}{7} \cos^7 x + C. \end{aligned}$$

Пример 3. $\int \cos^7 x dx$. Применим подстановку $t = \sin x$, тогда $dt = \cos x dx$,

$$\begin{aligned} \int \cos^7 x dx &= \int (1 - \sin^2 x)^3 \cos x dx = \int (1 - t^2)^3 dt = \\ &= \int (1 - 3t^2 + 3t^4 - t^6) dt = t - t^3 + \frac{3}{5} t^5 - \frac{1}{7} t^7 + C = \sin x - \sin^3 x + \\ &\quad + \frac{3}{5} \sin^5 x - \frac{1}{7} \sin^7 x + C. \end{aligned}$$

Примечание 2. Интегралы от нечетных положительных степеней $\sec x$ и $\operatorname{cosec} x$ могут быть взяты с помощью подстановок, указанных в теоремах 2 и 3. Но так как для этих интегралов указанные подстановки приводят к громоздким выкладкам, их целесообразнее интегрировать по частям.

Теорема 4. Если функция $R(\sin x, \cos x)$ имеет четный характер относительно $\sin x$ и $\cos x$, то интеграл $\int R(\sin x, \cos x) dx$ подстановкой $t = \operatorname{tg} x$ приводится к интегралу от рациональной функции.

Преобразуем подинтегральную функцию:

$$R(\sin x, \cos x) = R(\operatorname{tg} x \cos x, \cos x).$$

Правая часть последнего равенства, имея четный характер относительно $\cos x$, может содержать лишь четные степени $\cos x$, а так как $\cos^2 x = \frac{1}{\sec^2 x} = \frac{1}{1 + \operatorname{tg}^2 x}$, то вся подинтегральная функция является некоторой рациональной функцией от $\operatorname{tg} x$, т. е. $R(\sin x, \cos x) = r(\operatorname{tg} x)$.

Умножим и разделим $r(\operatorname{tg} x)$ на $\sec^2 x$ и затем применим подстановку $t = \operatorname{tg} x$, $dt = \sec^2 x dx$. Получим:

$$\begin{aligned} \int R(\sin x, \cos x) dx &= \int r(\operatorname{tg} x) dx = \\ &= \int \frac{r(\operatorname{tg} x)}{1 + \operatorname{tg}^2 x} \sec^2 x dx = \int r_1(t) dt. \end{aligned}$$

Пример 4. $\int \frac{dx}{\sin^2 x + 6 \sin x \cos x - 16 \cos^2 x}.$

Подинтегральная функция имеет четный характер относительно $\sin x$ и $\cos x$. Разделив все члены числителя и знаменателя на $\cos^2 x$, применим подстановку $t = \operatorname{tg} x$. Тогда $dt = \sec^2 x dx$,

$$\begin{aligned} \int \frac{dx}{\sin^2 x + 6 \sin x \cos x - 16 \cos^2 x} &= \int \frac{\sec^2 x dx}{\operatorname{tg}^2 x + 6 \operatorname{tg} x - 16} = \\ &= \int \frac{dt}{t^2 + 6t - 16} = \int \frac{dt}{(t+3)^2 - 25} = \frac{1}{10} \ln \left| \frac{\operatorname{tg} x + 3 - 5}{\operatorname{tg} x + 3 + 5} \right| + C = \\ &= \frac{1}{10} \ln \left| \frac{\operatorname{tg} x - 2}{\operatorname{tg} x + 8} \right| + C. \end{aligned}$$

Примечание 3. Интегралы от целых степеней (больших единицы) $\operatorname{tg} x$ и $\operatorname{ctg} x$ и от положительных четных степеней $\sec x$ и $\operatorname{cosec} x$ удобнее брать, используя соотношения: $\operatorname{tg}^2 x = \sec^2 x - 1$ и $\operatorname{ctg}^2 x = \operatorname{cosec}^2 x - 1$.

Пример 5. $\int \operatorname{tg}^4 x dx = \int \operatorname{tg}^2 x (\sec^2 x - 1) dx = \int \operatorname{tg}^2 x \sec^2 x dx -$
 $-\int \operatorname{tg}^2 x dx = \int \operatorname{tg}^2 x \sec^2 x dx - \int \sec^2 x dx + \int dx = \frac{1}{3} \operatorname{tg}^3 x - \operatorname{tg} x + x + C.$

Пример 6. $\int \operatorname{ctg}^3 x dx = \int \operatorname{ctg} x (\operatorname{cosec}^2 x - 1) dx = \int \operatorname{ctg} x \operatorname{cosec}^2 x dx -$
 $-\int \operatorname{ctg} x dx.$

$t = \operatorname{ctg} x$, $dt = -\operatorname{cosec}^2 x dx$; тогда

$$\int \operatorname{ctg}^3 x dx = -\int t dt - \int \operatorname{ctg} x dx = -\frac{1}{2} \operatorname{ctg}^2 x - \ln |\sin x| + C.$$

Пример 7. $\int \sec^8 x dx = \int (1 + \operatorname{tg}^2 x)^3 \sec^2 x dx.$

Применим подстановку $t = \operatorname{tg} x$; тогда $dt = \sec^2 x dx$,

$$\begin{aligned} \int \sec^8 x dx &= \int (1 + t^2)^3 dt = \int (1 + 3t^2 + 3t^4 + t^6) dt = \\ &= \operatorname{tg} x + \operatorname{tg}^3 x + \frac{3}{5} \operatorname{tg}^5 x + \frac{1}{7} \operatorname{tg}^7 x + C. \end{aligned}$$

Примечание 4. Из подинтегральных функций, удовлетворяющих условиям теоремы 4, на практике часто встречаются произведения неотрицательных четных степеней $\sin x$ и $\cos x$, т. е. функции вида $\sin^m x \cos^n x$, где m и n — четные неотрицательные числа. Интегрирование таких функций целесообразно производить, заменяя четные степени $\sin x$ и $\cos x$ синусами и косинусами кратных аргументов. Используем для этого известные формулы тригонометрии, которые запишем в следующей форме:

$$\cos^2 u = \frac{1}{2} (1 + \cos 2u), \quad (7)$$

$$\sin^2 u = \frac{1}{2} (1 - \cos 2u), \quad (8)$$

$$\sin^2 u \cos^2 u = \frac{1}{4} \sin^2 2u = \frac{1}{8} (1 - \cos 4u). \quad (9)$$

Пример 8. $\int \sin^2 7x dx$. Используем формулу (8):

$$\int \sin^2 7x dx = \frac{1}{2} \int (1 - \cos 14x) dx = \frac{1}{2} x - \frac{1}{28} \sin 14x + C.$$

Пример 9. $\int \cos^4 5x dx$. Используем формулу (7):

$$\begin{aligned} \int \cos^4 5x dx &= \frac{1}{4} \int (1 + \cos 10x)^2 dx = \frac{1}{4} \int (1 + 2 \cos 10x + \\ &+ \cos^2 10x) dx = \frac{1}{4} \int \left(1 + 2 \cos 10x + \frac{1}{2} + \frac{1}{2} \cos 20x \right) dx = \\ &= \frac{3}{8} x + \frac{1}{20} \sin 10x + \frac{1}{160} \sin 20x + C. \end{aligned}$$

Пример 10. $\int \sin^2 3x \cos^2 3x dx$. Используем формулу (9):

$$\int \sin^2 3x \cos^2 3x dx = \frac{1}{8} \int (1 - \cos 12x) dx = \frac{1}{8} x - \frac{1}{96} \sin 12x + C.$$

Интегрирование произведения синусов и косинусов различных аргументов. При интегрировании произведения синусов и косинусов различных аргументов: $\sin px \cos qx$, $\sin px \sin qx$, $\cos px \cos qx$ можно использовать тригонометрические формулы, приводящие сумму или разность синусов или косинусов к виду, удобному для логарифмирования.

Так, например, учитывая, что $\sin(\alpha + \beta)x + \sin(\alpha - \beta)x = 2 \sin \alpha x \cos \beta x$, можно представить интеграл $\int \sin px \cos qx dx$ в виде суммы двух интегралов:

$$\int \sin px \cos qx dx = \frac{1}{2} \int \sin(p+q)x dx + \frac{1}{2} \int \sin(p-q)x dx. \quad (10)$$

Аналогично:

$$\int \cos px \cos qx dx = \frac{1}{2} \int \cos(p+q)x dx + \frac{1}{2} \int \cos(p-q)x dx, \quad (11)$$

$$\int \sin px \sin qx dx = \frac{1}{2} \int \cos(p-q)x dx - \frac{1}{2} \int \cos(p+q)x dx. \quad (12)$$

Пример 11. $\int \sin 7x \sin 2x dx$. Применим формулу (12):

$$\begin{aligned} \int \sin 7x \sin 2x dx &= \frac{1}{2} \int \cos 5x dx - \frac{1}{2} \int \cos 9x dx = \\ &= \frac{1}{10} \sin 5x - \frac{1}{18} \sin 9x + C. \end{aligned}$$

Пример 12. $\int \sin 10x \cos 7x \cos 4x dx$. Применим сначала формулу (11), а затем формулу (10):

$$\begin{aligned} \int \sin 10x [\cos 7x \cos 4x] dx &= \frac{11}{2} \int \sin 10x \cos 11x dx + \\ &+ \frac{1}{2} \int \sin 10x \cos 3x dx = \frac{1}{4} \int \sin 21x dx + \frac{1}{4} \int \sin x dx + \\ &+ \frac{1}{4} \int \sin 13x dx + \frac{1}{4} \int \sin 7x dx = -\frac{1}{84} \cos 21x - \frac{1}{4} \cos x - \frac{1}{52} \cos 13x - \\ &- \frac{1}{28} \cos 7x + C. \end{aligned}$$

§ 11. Интегрирование иррациональных функций

I. Интегрирование функций, рациональных относительно аргумента и корня из дробно-линейной функции.

Теорема. Интеграл вида $\int R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right) dx$ (где n — натуральное число) подстановкой

$$\sqrt[n]{\frac{ax+b}{cx+d}} = t \quad (1)$$

приводится к интегралу от рациональной функции t или, как говорят, «рационализуется».

Пусть $\sqrt[n]{\frac{ax+b}{cx+d}} = t$, тогда $\frac{ax+b}{cx+d} = t^n$, $x = \frac{t^n - b}{a - ct^n}$, $dx = \left(\frac{t^n - b}{a - ct^n}\right)' dt$.

Следовательно,

$$\int R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right) dx = \int R\left(\frac{t^n - b}{a - ct^n}, t\right) \left(\frac{t^n - b}{a - ct^n}\right)' dt = \int r(t) dt.$$

Следствие. Интеграл вида $\int R(x, \sqrt[n]{ax+b}) dx$ подстановкой $\sqrt[n]{ax+b} = t$ рационализуется.

Примечание. Если подинтегральная функция является рациональной функцией относительно $\left(\frac{ax+b}{cx+d}\right)^{\frac{p}{q}}$, $\left(\frac{ax+b}{cx+d}\right)^{\frac{r}{s}}$ и т. д., где $p, q, r, s, \dots$ — целые числа, то интеграл рационализуется подстановкой $\frac{ax+b}{cx+d} = t^n$, где n — наименьшее кратное чисел $q, s, \dots$, и т. д.

Пример. $\int \frac{\sqrt[3]{x-1} + \sqrt[4]{x-1}}{(x-1)(1+\sqrt[6]{x-1})} dx$. Применим подстановку $\sqrt[12]{x-1} = t$,

тогда $x-1 = t^{12}$, $dx = 12t^{11}dt$,

$$\int \frac{\sqrt[3]{x-1} + \sqrt[4]{x-1}}{(x-1)(1+\sqrt[6]{x-1})} dx = 12 \int \frac{t^{11}(t^4 - t^3)}{t^{12}(1+t^2)} dt = 12 \int \left(t + 1 - \frac{t+1}{t^2+1} \right) dt = \\ = 6\sqrt[6]{x-1} + 12\sqrt[12]{x-1} - 6 \ln |\sqrt[6]{x-1} + 1| - 12 \arctg \sqrt[12]{x-1} + C.$$

II. Интегрирование биномиальных дифференциалов. Биномиальным дифференциалом называется выражение $x^m(a+bx^n)^p dx$, где m, n, p — рациональные числа.

Интеграл от биномиального дифференциала через элементарные функции выражается лишь в трех случаях: если в биномиальном дифференциале $x^m(a+bx^n)^p dx$

1) p — целое, то с помощью подстановки $\sqrt[\lambda]{x} = t$, где λ — общий знаменатель m и n , интеграл $\int x^m(a+bx^n)^p dx$ рационализуется;

2) $\frac{m+1}{n}$ — целое или равно 0, то подстановкой $\sqrt[s]{a+bx^n} = t$ (где s — знаменатель p) интеграл $\int x^m(a+bx^n)^p dx$ рационализуется;

3) $\frac{m+1}{n} + p$ — целое или равно 0, то подстановкой $\sqrt[s]{\frac{a+bx^n}{x^n}} = t$ (где s — знаменатель p) интеграл $\int x^m(a+bx^n)^p dx$ рационализуется¹⁾.

Академик П. Л. Чебышев (1821—1894) доказал, что во всех остальных случаях интеграл от биномиального дифференциала через элементарные функции не выражается.

III. Интегрирование функций, рациональных относительно аргумента и квадратного корня из квадратного трехчлена. Среди интегралов от иррациональных функций наибольшие практические применения имеют интегралы вида $\int R(x, \sqrt{ax^2+bx+c}) dx$. Рассмотрим несколько способов интегрирования этих функций.

А. Тригонометрические подстановки. Как было показано ранее, квадратный трехчлен ax^2+bx+c путем выделения полного квадрата и замены переменных может быть представлен в форме двухчлена вида $\pm u^2 \pm m^2$. Таким образом, достаточно ограничиться рассмотрением трех видов интегралов:

$$1) \int R(u, \sqrt{m^2 - u^2}) du; \quad 2) \int R(u, \sqrt{m^2 + u^2}) du; \quad 3) \int R(u, \sqrt{u^2 - m^2}) du.$$

Четвертая комбинация знаков $(-u^2 - m^2)$ приводит нас к подинтегральной функции, которая не существует в действительной области.

Теорема 1. Интеграл вида $\int R(u, \sqrt{m^2 - u^2}) du$ подстановкой $u = m \sin t$ ($u = m \cos t$) сводится к интегралу от рациональной функции относительно $\sin t$ и $\cos t$.

Для доказательства применим, например, подстановку $u = m \sin t$, $m > 0$, тогда $du = m \cos t dt$, $\sqrt{m^2 - u^2} = m \cos t$,

$$\int R(u, \sqrt{m^2 - u^2}) du = \int R(m \sin t, m \cos t) m \cos t dt.$$

¹⁾ В данном курсе доказательство этих трех положений не приводится.

Теорема 2. Интеграл вида $\int R(u, \sqrt{m^2 + u^2}) du$ подстановкой $u = m \operatorname{tg} t$ ($u = m \operatorname{ctg} t$) сводится к интегралу от рациональной функции относительно $\sin t$ и $\cos t$.

Для доказательства применим, например, подстановку $u = m \operatorname{tg} t$, $m > 0$, тогда $du = m \sec^2 t dt$, $\sqrt{m^2 + u^2} = m \sec t$,

$$\int R(u, \sqrt{m^2 + u^2}) du = \int R(m \operatorname{tg} t, m \sec t) m \sec^2 t dt.$$

Теорема 3. Интеграл вида $\int R(u, \sqrt{u^2 - m^2}) du$ подстановкой $u = m \sec t$ ($u = m \operatorname{cosec} t$) сводится к интегралу от рациональной функции относительно $\sin t$ и $\cos t$.

Для доказательства применим, например, подстановку $u = m \sec t$, $m > 0$, тогда $du = m \operatorname{tg} t \sec t dt$, $\sqrt{u^2 - m^2} = m \operatorname{tg} t$,

$$\int R(u, \sqrt{u^2 - m^2}) du = \int R(m \sec t, m \operatorname{tg} t) m \operatorname{tg} t \sec t dt.$$

Интегралы от рациональных функций относительно $\sin t$ и $\cos t$, как известно, могут быть выражены через интегралы от рациональных алгебраических функций.

Пример 1. $\int \sqrt{a^2 - x^2} dx$. Применим подстановку $x = a \sin t$, тогда $dx = a \cos t dt$, $\sqrt{a^2 - x^2} = a \cos t$,

$$\begin{aligned} \int \sqrt{a^2 - x^2} dx &= a^2 \int \cos^2 t dt = \frac{a^2}{2} \int (1 + \cos 2t) dt = \\ &= \frac{a^2}{2} t + \frac{a^2}{4} \sin 2t + C = \frac{a^2}{2} t + \frac{a^2}{2} \sin t \cos t + C. \end{aligned}$$

Выразим теперь t , $\sin t$, $\cos t$ через x : $\sin t = \frac{x}{a}$; $\cos t = \frac{\sqrt{a^2 - x^2}}{a}$,
 $t = \arcsin \frac{x}{a}$,

$$\int \sqrt{a^2 - x^2} dx = \frac{a^2}{2} \arcsin \frac{x}{a} + \frac{x}{a} \sqrt{a^2 - x^2} + C.$$

Пример 2. $\int \frac{dx}{x^4 \sqrt{a^2 + x^2}}$. Применим подстановку $x = a \operatorname{tg} t$, тогда $dx = a \sec^2 t dt$, $\sqrt{a^2 + x^2} = a \sec t$,

$$\begin{aligned} \int \frac{dx}{x^4 \sqrt{a^2 + x^2}} &= \int \frac{a \sec^2 t dt}{a^4 \operatorname{tg}^4 t \cdot a \sec t} = \frac{1}{a^4} \int \frac{\cos^3 t}{\sin^4 t} dt = \\ &= \frac{1}{a^4} \int \frac{(1 - \sin^2 t) d \sin t}{\sin^4 t} = -\frac{1}{3a^4 \sin^3 t} + \frac{1}{a^4 \sin t} + C. \end{aligned}$$

Выразим из соотношения $x = a \operatorname{tg} t$ функцию $\sin t$ через x : $\sin t = \frac{x}{\sqrt{a^2 + x^2}}$.

Тогда

$$\int \frac{dx}{x^4 \sqrt{a^2 + x^2}} = -\frac{(a^2 + x^2)^{3/2}}{3a^4 x^3} + \frac{\sqrt{a^2 + x^2}}{a^4 x} + C.$$

Пример 3. $\int \frac{dx}{x(x^2 - 4)^{5/2}}$. Применим подстановку $x = 2 \sec t$, тогда $dx = 2 \operatorname{tg} t \sec t dt$, $\sqrt{x^2 - 4} = 2 \operatorname{tg} t$.

$$\begin{aligned} \int \frac{dx}{x(x^2 - 4)^{5/2}} &= \int \frac{2 \operatorname{tg} t \sec t dt}{2 \sec t \cdot 2^5 \operatorname{tg}^5 t} = \frac{1}{32} \int \operatorname{ctg}^4 t dt = \\ &= \frac{1}{32} \int \operatorname{ctg}^2 t (\operatorname{cosec}^2 t - 1) dt = -\frac{1}{32} \int \operatorname{ctg}^2 t d \operatorname{ctg} t - \frac{1}{32} \int \operatorname{cosec}^2 t dt + \frac{1}{32} \int dt = \\ &= -\frac{1}{96} \operatorname{ctg}^3 t + \frac{1}{32} \operatorname{ctg} t + \frac{1}{32} t + C. \end{aligned}$$

Выразим $\operatorname{ctg} t$ и t через x : $t = \operatorname{arcsec} \frac{x}{2} = \arccos \frac{2}{x}$, $\operatorname{ctg} t = \frac{2}{\sqrt{x^2 - 4}}$.

$$\int \frac{dx}{x(x^2 - 4)^{5/2}} = -\frac{1}{12(x^2 - 4)^{3/2}} + \frac{1}{16\sqrt{x^2 - 4}} + \frac{1}{32} \arccos \frac{2}{x} + C.$$

Б. Подстановки Эйлера.

1. Если $a > 0$, то интеграл $\int R(x, \sqrt{ax^2 + bx + c}) dx$ рационализируется подстановкой $\sqrt{ax^2 + bx + c} = t \pm x \sqrt{a}$.

2. Если $a < 0$, $c > 0$, то интеграл $\int R(x, \sqrt{ax^2 + bx + c}) dx$ рационализируется подстановкой $\sqrt{ax^2 + bx + c} = tx \pm \sqrt{c}$.

3. Если $a < 0$, а подкоренное выражение раскладывается на действительные множители $a(x - x_1)(x - x_2)$, то интеграл $\int R(x, \sqrt{ax^2 + bx + c}) dx$ рационализируется подстановкой $\sqrt{ax^2 + bx + c} = t(x - x_1)$.

Следует заметить, что подстановки Эйлера имеют скорее теоретическое, чем практическое значение. Даже при несложных интегралах подстановки Эйлера нередко приводят к громоздким выкладкам. Поэтому в данном курсе эти подстановки подробно не рассматриваются.

В. Метод неопределенных коэффициентов. Вычисление интегралов вида $\int R(x, \sqrt{ax^2 + bx + c}) dx$ часто сводится к вычислению интегралов следующих трех типов:

$$\text{I. } \int \frac{P(x) dx}{\sqrt{ax^2 + bx + c}}, \quad \text{II. } \int P(x) \sqrt{ax^2 + bx + c} dx, \\ \text{III. } \int \frac{dx}{(x - a)^n \sqrt{ax^2 + bx + c}},$$

где $P(x)$ — многочлен.

Покажем, что интегралы типов II и III могут быть сведены к интегралу типа I.

Для преобразования интеграла типа II достаточно числитель и знаменатель умножить на радикал $\sqrt{ax^2 + bx + c}$. Действительно,

$$\int P(x) \sqrt{ax^2 + bx + c} dx = \int \frac{P(x)(ax^2 + bx + c)}{\sqrt{ax^2 + bx + c}} dx = \int \frac{P_1(x) dx}{\sqrt{ax^2 + bx + c}}.$$

Для приведения интеграла типа III к интегралу типа I применяют так называемую «обратную» подстановку $x - a = \frac{1}{v}$, тогда $dx = -\frac{dv}{v^2}$, $x = \frac{1 + av}{v}$,

$$\int \frac{dx}{(x - a)^n \sqrt{ax^2 + bx + c}} = - \int \frac{v^n dv}{v^2 \sqrt{\frac{a(1 + av)^2}{v^2} + \frac{b(1 + av)}{v} + c}} = \\ = \int \frac{P_2(v) dv}{\sqrt{a_1 v^3 + b_1 v + c_1}},$$

где a_1, b_1, c_1 — коэффициенты трехчлена, полученные после приведения подобных членов, $P_2(v)$ — многочлен.

Следовательно, интегралы типов II и III легко приводятся к интегралу типа I. Рассматривая интегралы вида $\int \frac{P(x) dx}{\sqrt{ax^2 + bx + c}}$, где $P(x)$ — многочлен, можно доказать, что

$$\int \frac{P(x) dx}{\sqrt{ax^2 + bx + c}} = Q(x) \sqrt{ax^2 + bx + c} + \lambda \int \frac{dx}{\sqrt{ax^2 + bx + c}}, \quad (2)$$

где $Q(x)$ — некоторый многочлен, степень которого ниже степени многочлена $P(x)$; λ — некоторая постоянная.

Практически при вычислении указанного вида интегралов применяется *метод неопределенных коэффициентов*. Для этого записывают интеграл по формуле (2), в которой λ — неопределенный коэффициент и $Q(x)$ — многочлен с неопределенными коэффициентами степени на единицу ниже степени $P(x)$.

Для нахождения неопределенных коэффициентов дифференцируют обе части равенства (2), затем умножают на $\sqrt{ax^2+bx+c}$ и, сравнивая коэффициенты при одинаковых степенях x , определяют λ и коэффициенты многочлена $Q(x)$.

Заметим, что формулу (2) обычно применяют тогда, когда степень многочлена $P(x)$ не меньше двух. Если $P(x)$ — линейная функция, то проще пользоваться методом, указанным в разделе интегрирования элементарных дробей.

Пример 1. $\int \frac{3x^3-7x^2+1}{\sqrt{x^2-2x+5}} dx$. Применим формулу (2):

$$\int \frac{3x^3-7x^2+1}{\sqrt{x^2-2x+5}} dx = (Ax^2+Bx+C) \sqrt{x^2-2x+5} + \lambda \int \frac{dx}{\sqrt{x^2-2x+5}}. \quad (3)$$

Продифференцируем равенство (3), затем умножим на $\sqrt{x^2-2x+5}$ и, сравнивая коэффициенты при одинаковых степенях x , определим коэффициенты A, B, C, λ :

$$\begin{aligned} \frac{3x^3-7x^2+1}{\sqrt{x^2-2x+5}} &= (2Ax+B) \sqrt{x^2-2x+5} + \frac{(Ax^2+Bx+C)}{\sqrt{x^2-2x+5}}(x-1) + \frac{\lambda}{\sqrt{x^2-2x+5}}, \\ 3x^3-7x^2+1 &= 2Ax^3+Bx^2-4Ax^2-2Bx+10Ax+5B+Ax^3+Bx^2+Cx- \\ &\quad -Ax^2-Bx-C+\lambda, \\ A=1, \quad B=-1, \quad C=-13, \quad \lambda &=-7. \end{aligned}$$

Подставим найденные значения в равенство (3) и возьмем интеграл, находящийся в правой части:

$$\begin{aligned} \int \frac{3x^3-7x^2+1}{\sqrt{x^2-2x+5}} dx &= (x^2-x-13) \sqrt{x^2-2x+5} - \\ &\quad - 7 \ln |x-1+\sqrt{x^2-2x+5}| + C. \end{aligned}$$

Пример 2. $\int (4x^2-6x) \sqrt{x^2+3} dx$.

Умножим числитель и знаменатель подинтегрального выражения на $\sqrt{x^2+3}$, а затем будем брать интеграл так же, как и в примере 1.

$$\begin{aligned} \int \frac{(4x^2-6x)(x^2+3)}{\sqrt{x^2+3}} dx &= \\ &= (Ax^3+Bx^2+Cx+D) \sqrt{x^2+3} + \lambda \int \frac{dx}{\sqrt{x^2+3}}, \\ \frac{4x^4-6x^3+12x^2-18x}{\sqrt{x^2+3}} &= (3Ax^2+2Bx+C) \sqrt{x^2+3} + \frac{(Ax^3+Bx^2+Cx+D)x}{\sqrt{x^2+3}} + \\ &\quad + \frac{\lambda}{\sqrt{x^2+3}}, \end{aligned}$$

$$4x^4-6x^3+12x^2-18x = 3Ax^4+2Bx^3+Cx^2+9Ax^2+6Bx+3C+Ax^4+Bx^3+ \\ + Cx^2+Dx+\lambda,$$

$$A=1; \quad B=-2; \quad C=\frac{3}{2}; \quad D=-6; \quad \lambda=-\frac{9}{2},$$

$$\begin{aligned} \int (4x^2-6x) \sqrt{x^2+3} dx &= \left(x^3-2x^3+\frac{3}{2}x-6\right) \sqrt{x^2+3} - \\ &\quad - \frac{9}{2} \ln |x+\sqrt{x^2+3}| + C. \end{aligned}$$

Пример 3. $\int \frac{4dx}{x^3 \sqrt{4x^2-1}}$.

Применив обратную подстановку $x = \frac{1}{v}$, $dx = -\frac{dv}{v^2}$, приведем данный интеграл к интегралу типа I:

$$\begin{aligned}\int \frac{4dx}{x^3 \sqrt{4x^2-1}} &= - \int \frac{4v^3 dv}{v^2 \sqrt{\frac{4}{v^2}-1}} = - \int \frac{4v^2 dv}{\sqrt{4-v^2}}, \\ &- \int \frac{4v^2 dv}{\sqrt{4-v^2}} = (Av + B) \sqrt{4-v^2} + \lambda \int \frac{dv}{\sqrt{4-v^2}}, \\ \frac{-4v^2}{\sqrt{4-v^2}} &= A \sqrt{4-v^2} + \frac{(Av+B)(-v)}{\sqrt{4-v^2}} + \frac{\lambda}{\sqrt{4-v^2}}, \\ -4v^2 &= 4A - Av^2 - Av^2 - Bv + \lambda, \\ A &= 2, \quad B = 0, \quad \lambda = -8. \\ \int \frac{4dx}{x^3 \sqrt{4x^2-1}} &= 2v \sqrt{4-v^2} - 8 \arcsin \frac{v}{2} + C = \\ &= \frac{2 \sqrt{4x^2-1}}{x^3} - 8 \arcsin \frac{1}{2x} + C.\end{aligned}$$

§ 12. Некоторые типы интегралов, берущихся методом интегрирования по частям

В § 4 рассматривались некоторые типы интегралов, берущихся по частям. После изучения основных приемов интегрирования различных элементарных функций область применения метода интегрирования по частям можно расширить.

Рассмотрим некоторые интегралы, берущиеся методом интегрирования по частям или сводящиеся после применения этого метода к интегралам, берущимся другими известными методами.

Пусть $P(x)$ и $Q(x)$ — многочлены.

$$1. \quad \int P(x) \ln |Q(x)| dx. \quad (1)$$

Обозначим $u = \ln |Q(x)|$, $dv = P(x) dx$, тогда $du = \frac{Q'(x)}{Q(x)}$, $v = P_1(x)$, где $P_1(x)$ — многочлен, степень которого на единицу выше степени многочлена $P(x)$.

$$\int P(x) \ln |Q(x)| dx = P_1(x) \ln |Q(x)| - \int \frac{P_1(x) Q'(x)}{Q(x)} dx. \quad (2)$$

Интеграл в правой части равенства (2) является интегралом от рациональной функции x и, следовательно, всегда может быть взят.

$$2. \quad \int P(x) \operatorname{arctg} [Q(x)] dx. \quad (3)$$

Обозначим $u = \operatorname{arctg} [Q(x)]$, $dv = P(x) dx$, тогда $du = \frac{Q'(x) dx}{1+[Q(x)]^2}$, $v = P_1(x)$,

$$\int P(x) \operatorname{arctg} [Q(x)] dx = P_1(x) \operatorname{arctg} [Q(x)] - \int \frac{P_1(x) Q'(x) dx}{1+[Q(x)]^2}. \quad (4)$$

В правой части равенства (4) находится интеграл от рациональной функции x , следовательно, его всегда можно найти.

$$3. \quad \int P(x) \arcsin x dx. \quad (5)$$

Обозначим $u = \arcsin x$, $dv = P(x) dx$, тогда $du = \frac{dx}{\sqrt{1-x^2}}$, $v = P_1(x)$,

$$\int P(x) \arcsin x dx = P_1(x) \arcsin x - \int \frac{P_1(x) dx}{\sqrt{1-x^2}}. \quad (6)$$

В правой части равенства (6) получается интеграл от иррациональной функции, берущийся, например, с помощью тригонометрической под-

становки. Следовательно, для интеграла вида (3) метод интегрирования по частям приводит к цели.

$$4. \quad \int e^{\alpha x} [P(x) \cos \beta x + Q(x) \sin \beta x] dx, \quad (7)$$

где $P(x)$ и $Q(x)$ — многочлены.

Можно показать, что интеграл вида (7) также берется методом интегрирования по частям. При интегрировании в правой части получается выражение, «подобное» подинтегральной функции:

$$\int e^{\alpha x} [P(x) \cos \beta x + Q(x) \sin \beta x] dx = e^{\alpha x} [M(x) \cos \beta x + N(x) \sin \beta x], \quad (8)$$

где $M(x)$ и $N(x)$ — многочлены степени не выше наибольшей из степеней $P(x)$ и $Q(x)$. Записав в формуле (8) вместо $M(x)$ и $N(x)$ многочлены с неопределенными коэффициентами степени, равной большей из степеней многочленов $P(x)$ и $Q(x)$, дифференцируя левую и правую части получившегося равенства, методом сравнения коэффициентов можно найти коэффициенты многочленов $M(x)$ и $N(x)$. Легко применяется этот прием для частного случая интеграла (7), а именно, для интеграла вида $\int P(x) e^{\alpha x} dx$.

Пример 1. $\int e^{2x} [(17x - 12) \cos 3x + (7x + 5) \sin 3x] dx$. На основании равенства (8) запишем:

$$\begin{aligned} \int e^{2x} [(17x - 12) \cos 3x + (7x + 5) \sin 3x] dx = \\ = e^{2x} [(Ax + B) \cos 3x + (Cx + D) \sin 3x]. \end{aligned} \quad (9)$$

Определим A, B, C, D :

$$\begin{aligned} e^{2x} [(17x - 12) \cos 3x + (7x + 5) \sin 3x] = e^{2x} [(2Ax + 2B) \cos 3x + \\ + (2Cx + 2D) \sin 3x + A \cos 3x + C \sin 3x + (3Cx + 3D) \cos 3x + \\ + (-3Ax - 3B) \sin 3x], \end{aligned}$$

откуда $A = 1, B = -2, C = 5, D = -3$.

Следовательно,

$$\begin{aligned} \int e^{2x} [(17x - 12) \cos 3x + (7x + 5) \sin 3x] dx = \\ = e^{2x} [(x - 2) \cos 3x + (5x - 3) \sin 3x] + C. \end{aligned}$$

Пример 2. $I = \int e^{5x} (15x^4 + 12x^3 - 10x^2 + 31x - 18) dx$. Применим метод неопределенных коэффициентов:

$$\int e^{5x} (15x^4 + 12x^3 - 10x^2 + 31x - 18) dx = e^{5x} (Ax^4 + Bx^3 + Cx^2 + Dx + E),$$

дифференцируем обе части равенства:

$$\begin{aligned} e^{5x} (15x^4 + 12x^3 - 10x^2 + 31x - 18) = \\ = e^{5x} (5Ax^4 + 5Bx^3 + 5Cx^2 + 5Dx + 5E + 4Ax^3 + 3Bx^2 + 2Cx + D), \end{aligned}$$

откуда $A = 3, B = 0, C = -2, D = 7, E = -5$, следовательно, $I = e^{5x} (3x^4 - 2x^2 + 7x - 5) + C$.

§ 13. Примеры интегралов, не выражающихся через элементарные функции

В предшествующих параграфах рассматривались классы элементарных функций, интегралы от которых выражаются через элементарные функции.

На практике часто встречаются и такие элементарные функции, интегралы от которых не выражаются через элементарные функции. Интегралы вида $\int R(x, \sqrt{P(x)}) dx$, где $P(x)$ — многочлен степени выше второй, в общем случае не выражаются через элементарные функции. При этом если $P(x)$ — многочлен третьей или четвертой степени, то они называются *эллиптическими*, если же степени выше четвертой, то — *ультраэллипти-*

ческими. В некоторых частных случаях интегралы вида $\int R(x, \sqrt{P(x)}) dx$ могут выражаться и через элементарные функции, они называются *псевдоэллиптическими*.

Из интегралов от трансцендентных функций, не выражающихся через элементарные функции и имеющих большое прикладное значение, следует указать *интеграл Пуассона*: $\int e^{-x^2} dx$, *интегралы Френеля*: $\int \sin x^2 dx$, $\int \cos x^2 dx$, *интегральный логарифм*: $\int \frac{dx}{\ln x}$ и сводящийся к нему интеграл $\int \frac{e^x}{x} dx$, *интегральный синус*: $\int \frac{\sin x}{x} dx$ и *интегральный косинус*: $\int \frac{\cos x}{x} dx$.

ГЛАВА VI

ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

Раздел 1

ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ И ЕГО СВОЙСТВА

§ 1. Задачи, приводящие к понятию определенного интеграла

1. **Задача о площади.** В плоскости XOY дана фигура $ABCD$, ограниченная отрезком $[a, b]$ оси OX , прямыми $x=a$ и $x=b$ и кривой $y=f(x)$, где $f(x)$ — непрерывная, неотрицательная на отрезке $[a, b]$ функция (рис. 243). Будем называть такую фигуру *криволинейной трапецией* с основанием $[a, b]$. В частности, точка D может совпасть с точкой A , а точка C — с точкой B . Требуется определить площадь S этой криволинейной трапеции.

Для решения поставленной задачи можно поступить следующим образом. Разобьем отрезок $[a, b]$ оси OX на n частей точками $x_0=a$,

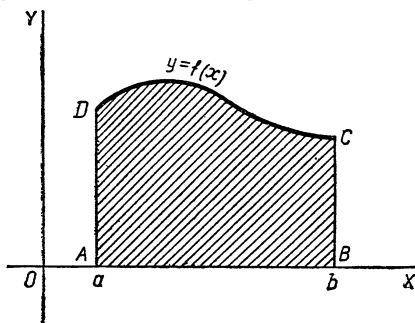


Рис. 243.



Рис. 244.

$x_1, x_2, \dots, x_n=b$ ($x_0 < x_1 < x_2 < \dots < x_n$). Через точки деления проведем вертикальные прямые. При этом криволинейная трапеция $ABCD$ разобьется на n криволинейных трапеций с основаниями $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$. Взяв на каждом отрезке $[x_{i-1}, x_i]$, $i=1, 2, \dots, n$, произвольно точку ξ_i ($x_{i-1} \leq \xi_i \leq x_i$), построим прямоугольник с основанием $[x_{i-1}, x_i]$ и высотой $f(\xi_i)$. Его площадь равна произведению $f(\xi_i) \cdot (x_i - x_{i-1})$.

Рассмотрим ступенчатую фигуру, составленную из прямоугольников с основаниями $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ (рис. 244). Ее площадь, равную сумме $f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \dots + f(\xi_n)(x_n - x_{n-1})$, будем считать приближенно равной площади S криволинейной трапеции $ABCD$:

$$S \approx f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \dots + f(\xi_n)(x_n - x_{n-1})$$

или короче:

$$S = \sum_{i=1}^n f(\xi_i) (x_i - x_{i-1})^1).$$

Чем мельче отрезки деления, тем точнее эта ступенчатая фигура «изображает» криволинейную трапецию $ABCD$.

За площадь криволинейной трапеции естественно принять предел²⁾, к которому стремятся площади построенных указанным образом ступенчатых фигур при неограниченном увеличении числа отрезков деления и стремлении к нулю длин отрезков деления.

Таким образом,

$$S = \lim \sum_{i=1}^n f(\xi_i) (x_i - x_{i-1})$$

или, если обозначить $x_i - x_{i-1} = \Delta x_i$,

$$S = \lim \sum_{i=1}^n f(\xi_i) \Delta x_i.$$

2. Задача о работе переменной силы. Материальная точка перемещается под действием силы F , направленной вдоль оси OX и имеющей переменную величину: $|F| = f(x)$. Требуется определить работу A , совершаемую силой F по перемещению единицы массы вдоль оси OX из точки $x=a$ в точку $x=b$ ($a < b$). Функцию $f(x)$ предполагаем непрерывной на отрезке $[a, b]$.

Если бы величина силы F была постоянна, то работа (по определению) равнялась бы произведению величины силы на длину пути перемещения. В случае силы переменной величины работу можно определить следующим образом.

Разобьем отрезок $[a, b]$ точками $x_0 = a, x_1, x_2, \dots, x_n = b$ ($x_0 < x_1 < x_2 < \dots < x_n$) на n отрезков: $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ и выберем на каждом отрезке $[x_{i-1}, x_i]$, $i = 1, 2, \dots, n$, произвольно точку ξ_i . Сила, действующая на отрезке $[x_{i-1}, x_i]$, меняется от точки к точке. Но если длина отрезка мала, ее величина в точках отрезка $[x_{i-1}, x_i]$ мало отличается от ее значения в точке ξ_i , так как функция $f(x)$ непрерывна. Поэтому работу A_i , совершаемую силой F на отрезке $[x_{i-1}, x_i]$, естественно считать приближенно равной работе, совершаемой на том же участке постоянной силой величины $f(\xi_i)$:

$$A_i \approx f(\xi_i) \Delta x_i.$$

Проведя такие рассуждения для каждого отрезка деления, получаем приближенное значение работы A силы F на всем отрезке $[a, b]$:

$$A \approx f(\xi_1) \Delta x_1 + f(\xi_2) \Delta x_2 + \dots + f(\xi_n) \Delta x_n = \sum_{i=1}^n f(\xi_i) \Delta x_i.$$

Точное значение работы A получаем в пределе при неограниченном увеличении числа отрезков деления и стремлении к нулю длин отрезков деления. Таким образом,

$$A = \lim \sum_{i=1}^n f(\xi_i) \Delta x_i.$$

¹⁾ Символ $\sum_{i=1}^n f(\xi_i) (x_i - x_{i-1})$ обозначает сумму всех выражений вида $f(\xi_i) (x_i - x_{i-1})$, получаемых при $i = 1, 2, \dots, n$.

²⁾ Точное определение предела указанного вида дается в следующем параграфе.

§ 2. Интегральная сумма. Определенный интеграл

Решение рассмотренных выше задач сводилось к отысканию пределов сумм специального вида при неограниченном увеличении числа отрезков деления и стягивании каждого из них в точку. Так как к нахождению такого рода пределов приводит огромное количество математических и прикладных задач, естественно изучить эти суммы и их пределы, абстрагируясь от их конкретного содержания.

Пусть на отрезке $[a, b]$ оси OX задана произвольная функция $f(x)$ ¹⁾. Разобьем отрезок $[a, b]$ на n частей точками $x_0 = a, x_1, x_2, \dots, x_n = b$ ($x_0 < x_1 < x_2 < \dots < x_n$). На каждом из отрезков деления $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$, внутри или на концах, выберем по точке. Обозначим эти точки соответственно $\xi_1, \xi_2, \dots, \xi_n$ ($x_{i-1} \leq \xi_i \leq x_i, i = 1, 2, \dots, n$). Составим произведения значений функции $f(x)$ в выбранных точках на длины соответствующих отрезков деления:

$$f(\xi_1)(x_1 - x_0), f(\xi_2)(x_2 - x_1), \dots, f(\xi_n)(x_n - x_{n-1}).$$

Сумма всех таких произведений

$$f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \dots + f(\xi_n)(x_n - x_{n-1})$$

или короче:

$$\sum_{i=1}^n f(\xi_i) \Delta x_i \quad (\Delta x_i = x_i - x_{i-1})$$

называется *интегральной суммой* для функции $f(x)$ на отрезке $[a, b]$. Она зависит от способа разбиения отрезка $[a, b]$ на части $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ и от выбора точек $\xi_1, \xi_2, \dots, \xi_n$ на этих частях. Таким образом, для функции $f(x)$ на отрезке $[a, b]$ можно составить бесчисленное множество интегральных сумм.

Назовем *шагом разбиения* отрезка $[a, b]$ на части наибольшую из длин отрезков деления. Процесс, состоящий в неограниченном увеличении числа n отрезков деления и стремлении к нулю длин отрезков деления, полностью характеризуется словами: «шаг разбиения отрезка $[a, b]$ стремится к нулю». При таком процессе как-то будут изменяться интегральные суммы.

Будем говорить, что интегральные суммы функции $f(x)$ при стремлении шага разбиения отрезка $[a, b]$ к нулю имеют *пределом* число I , если для любого числа $\varepsilon > 0$ существует число $\delta > 0$, такое, что все интегральные суммы с шагом разбиения, меньшим δ , удовлетворяют неравенству

$$\left| I - \sum_{i=1}^n f(\xi_i) \Delta x_i \right| < \varepsilon,$$

т. е. отличаются от числа I по абсолютной величине меньше чем на ε .

Число I , если оно существует, очевидно, зависит только от вида функции $f(x)$ и отрезка $[a, b]$ и не зависит от способов разбиения отрезка $[a, b]$ на части и выбора точек ξ_i на этих частях. Это число I и называется *определенным интегралом* от функции $f(x)$ на отрезке $[a, b]$ и обозначается символом

$$\int_a^b f(x) dx$$

(«интеграл от a до b эф от икс де икс»).

¹⁾ Заметим, что никаких ограничений на знак функции не накладывается.

Таким образом, определенный интеграл от функции $f(x)$ на отрезке $[a, b]$ есть число, равное пределу интегральных сумм при стремлении к нулю шага разбиения отрезка $[a, b]$:

$$\int_a^b f(x) dx = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (\lambda — \text{шаг разбиения отрезка } [a, b]).$$

При этом числа a и b называются соответственно *нижним и верхним пределами интегрирования*, отрезок $[a, b]$ — *промежутком интегрирования*, функция $f(x)$ называется *подинтегральной функцией*, выражение $f(x) dx$ — *подинтегральным выражением*, x — *переменной интегрирования*.

§ 3. Условия существования определенного интеграла

Функция $f(x)$, для которой на отрезке $[a, b]$ существует определенный интеграл, называется *интегрируемой* на этом отрезке.

Естественно возникает вопрос: при каких условиях функция $f(x)$, определенная на отрезке $[a, b]$, интегрируема на этом отрезке.

Заметим прежде всего, что всякая интегрируемая функция должна быть ограничена.

Действительно, если функция $f(x)$ не ограничена на отрезке $[a, b]$, то для любого разбиения отрезка $[a, b]$ на части $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ найдется хоть один отрезок деления $[x_{k-1}, x_k]$, на котором функция будет неограниченной. Зафиксировав точки $\xi_1, \dots, \xi_{k-1}, \xi_{k+1}, \dots, \xi_n$, составим сумму

$$\Sigma' = f(\xi_1) \Delta x_1 + \dots + f(\xi_{k-1}) \Delta x_{k-1} + f(\xi_{k+1}) \Delta x_{k+1} + \dots + f(\xi_n) \Delta x_n.$$

Эта сумма не является интегральной, так как в ней отсутствует слагаемое, соответствующее отрезку $[x_{k-1}, x_k]$.

В силу неограниченности функции $f(x)$ на отрезке $[x_{k-1}, x_k]$ соответствующую ему точку ξ_k можно выбрать так, что абсолютная величина произведения $f(\xi_k) \Delta x_k$ будет больше любого наперед заданного числа. В частности, точку ξ_k можно выбрать так, что сумма

$$\Sigma' + f(\xi_k) \Delta x_k,$$

т. е. интегральная сумма

$$\sum_{i=1}^n f(\xi_i) \Delta x_i,$$

будет по абсолютной величине больше любого наперед заданного числа.

Так как для любого разбиения отрезка $[a, b]$ на части найдутся интегральные суммы, сколь угодно большие по абсолютной величине, то не существует (конечного) предела интегральных сумм при стремлении шага разбиения к нулю.

Таким образом, *необходимым условием* интегрируемости функции на отрезке $[a, b]$ является ее ограниченность на этом отрезке.

Рассмотрим теперь *достаточные условия* интегрируемости.

Теорема 1. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на этом отрезке.

Для доказательства этой теоремы введем понятия нижних и верхних интегральных сумм и изучим их свойства.

Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$. Разобьем этот отрезок точками $x_0 = a < x_1 < x_2 < \dots < x_n = b$ на n отрезков: $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$. Обозначим полученное разбиение через T .

На каждом отрезке $[x_{i-1}, x_i]$, $i = 1, 2, \dots, n$, непрерывная функция $f(x)$ в каких-то точках принимает свое наименьшее (m_i) и наибольшее (M_i) значения.

значения на этом отрезке. Выбирая эти точки за точки ξ_i при составлении интегральных сумм, получим интегральные суммы

$$s = \sum_{i=1}^n m_i \Delta x_i \quad \text{и} \quad S = \sum_{i=1}^n M_i \Delta x_i.$$

Они называются соответственно *нижней* и *верхней интегральными суммами* функции $f(x)$ для данного разбиения T отрезка $[a, b]$. Так как для каждой точки x отрезка $[x_{i-1}, x_i]$ справедливы неравенства

$$m_i \leq f(x) \leq M_i,$$

то любая другая интегральная сумма $\sigma = \sum_{i=1}^n f(\xi_i) \Delta x_i$ для того же разбиения T заключена между s и S :

$$s \leq \sigma \leq S. \quad (1)$$

Если функция $f(x)$ неотрицательна на отрезке $[a, b]$, нижняя и верхняя интегральные суммы для данного разбиения T имеют простой геометрический смысл. Рассмотрим на плоскости XOY криволинейную трапецию, ограниченную отрезком $[a, b]$ оси OX , прямыми $x=a$ и $x=b$ и кривой $y=f(x)$. Произведение $m_i \Delta x_i$ равно площади прямоугольника с основанием $[x_{i-1}, x_i]$ и высотой m_i . Этот прямоугольник «вписан» в криволинейную трапецию с тем же

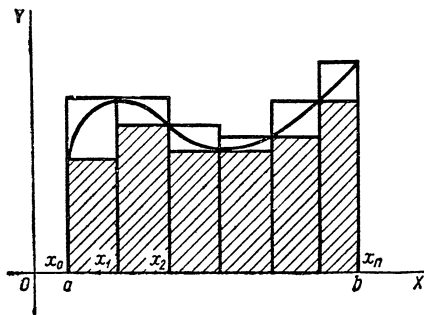


Рис. 245.

основанием. Сумма $s = \sum_{i=1}^n m_i \Delta x_i$ представляет собой площадь составленной из таких прямоугольников ступенчатой фигуры, «вписанной» в данную криволинейную трапецию

с основанием $[a, b]$. Аналогично заключаем, что сумма $S = \sum_{i=1}^n M_i \Delta x_i$ представляет собой площадь ступенчатой фигуры, «описанной» около криволинейной трапеции с основанием $[a, b]$ (см. рис. 245).

Любая другая интегральная сумма $\sigma = \sum_{i=1}^n f(\xi_i) \Delta x_i$ для того же разбиения T равна площади «промежуточной» ступенчатой фигуры, составленной из прямоугольников с теми же основаниями.

Нижние и верхние интегральные суммы функции $f(x)$ обладают следующими свойствами.

1. Множество значений всех верхних и нижних сумм для любых разбиений ограничено.

Действительно, если обозначить через m и M соответственно наименьшее и наибольшее значения функции $f(x)$ на отрезке $[a, b]$, то для любого разбиения T будем иметь:

$$m \sum_{i=1}^n \Delta x_i \leq \sum_{i=1}^n m_i \Delta x_i \leq \sum_{i=1}^n M_i \Delta x_i \leq M \sum_{i=1}^n \Delta x_i,$$

т. е.

$$m(b-a) \leq s \leq S \leq M(b-a).$$

2. От добавления к данному разбиению T новых точек деления нижняя сумма не убывает, а верхняя не возрастает.

Доказательство. Пусть к разбиению T добавлена одна новая точка деления (добавление нескольких точек деления можно произвести, добавляя их по одной). Предположим, что эта новая точка x'_i попала на отрезок $[x_{i-1}, x_i]$ (рис. 246). Обозначим через s и s' — нижние, а через S и S' — верхние интегральные суммы для данного разбиения T и полученного из него добавлением точки x'_i разбиения T' .

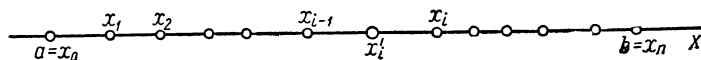


Рис. 246.

Рассмотрим суммы s и s' . Они отличаются между собой лишь членами, соответствующими отрезку $[x_{i-1}, x_i]$: в сумму s входит слагаемое $m_i \Delta x_i$ ($\Delta x_i = x_i - x_{i-1}$), а в сумму s' вместо него входит два слагаемых: $m'_i(x'_i - x_{i-1}) + m''_i(x_i - x'_i)$, где m'_i и m''_i — наименьшие значения функции $f(x)$ на отрезках $[x_{i-1}, x'_i]$ и $[x'_i, x_i]$. Так как $m'_i \geq m_i$, $m''_i \geq m_i$, заключаем, что

$$m'_i(x'_i - x_{i-1}) + m''_i(x_i - x'_i) \geq m_i(x'_i - x_{i-1} + x_i - x'_i) = m_i \Delta x_i,$$

а значит, $s' \geq s$.

Аналогично доказывается, что $S' \leq S$.

3. Нижняя сумма s любого разбиения T не превосходит верхней суммы S' любого другого разбиения T' .

Для доказательства этого утверждения рассмотрим разбиение T'' , состоящее из всех точек, входящих в разбиения T и T' . Это разбиение T'' можно получить как из разбиения T , так и из разбиения T' добавлением конечного числа точек деления.

Обозначая через s'' и S'' соответственно нижнюю и верхнюю интегральные суммы для разбиения T'' , на основании свойства 2 получаем:

$$s \leq s'' \leq S'' \leq S',$$

т. е. $s \leq S'$.

Докажем теперь сформулированную выше теорему об интегрируемости функции $f(x)$, непрерывной на отрезке $[a, b]$.

Рассмотрим последовательность разбиений $T_1, T_2, \dots, T_n, \dots$ отрезка $[a, b]$, удовлетворяющую двум условиям: 1) каждое разбиение этой последовательности получено из предыдущего добавлением конечного числа точек деления, 2) шаг разбиения T_n (обозначим его λ_n) стремится к нулю при $n \rightarrow \infty$.

Составив для каждого разбиения T_n нижнюю и верхнюю суммы s_n и S_n , получим две последовательности интегральных сумм:

$$s_1, s_2, \dots, s_n, \dots, \\ S_1, S_2, \dots, S_n, \dots$$

Из доказанных выше свойств следует, что последовательность нижних интегральных сумм монотонно возрастает (не убывает) и ограничена:

$$m(b-a) \leq s_1 \leq s_2 \leq \dots \leq s_n \leq \dots \leq M(b-a),$$

а последовательность верхних интегральных сумм монотонно убывает (не возрастает) и также ограничена:

$$m(b-a) \leq \dots \leq S_n \leq \dots \leq S_2 \leq S_1 \leq M(b-a).$$

Всякая монотонная ограниченная последовательность имеет конечный предел. Обозначим

$$\lim_{n \rightarrow \infty} s_n = I, \quad \lim_{n \rightarrow \infty} S_n = \bar{I}.$$

Для любого разбиения T_n

$$s_n \leq \underline{I}, \quad \bar{I} \leq S_n. \quad (2)$$

Переходя к пределу в неравенстве

$$s_i \leq S_k$$

(см. свойство 3), справедливым при любых i и k , получаем неравенство

$$\underline{I} \leq \bar{I}.$$

Сопоставляя его с неравенствами (2), заключаем, что для любого разбиения T_n

$$s_n \leq \underline{I} \leq \bar{I} \leq S_n. \quad (3)$$

Покажем, что $\underline{I} = \bar{I}$. Для этого, как следует из неравенств (3), достаточно доказать, что

$$\lim_{n \rightarrow \infty} (S_n - s_n) = 0.$$

Пусть $\varepsilon > 0$ — любое число. Функция $f(x)$, непрерывная на отрезке $[a, b]$, равномерно непрерывна на этом отрезке. Значит, для числа $\varepsilon_1 = \frac{\varepsilon}{b-a}$ существует число $\delta > 0$, такое, что $|f(x') - f(x'')| < \varepsilon_1$ для любых двух точек x' и x'' отрезка $[a, b]$, находящихся на расстоянии, меньшем δ .

Так как $\lim_{n \rightarrow \infty} \lambda_n = 0$, то существует номер N , начиная с которого шаг разбиения λ_n меньше δ .

При $n > N$

$$S_n - s_n = \sum_{i=1}^k (M_i - m_i) \Delta x_i < \varepsilon_1 \sum_{i=1}^k \Delta x_i = \varepsilon_1 (b - a) = \varepsilon,$$

откуда следует, что

$$\lim_{n \rightarrow \infty} (S_n - s_n) = 0,$$

а значит,

$$\underline{I} = \bar{I} = I.$$

Покажем теперь, что по любой другой последовательности разбиений

$$T'_1, T'_2, \dots, T'_n, \dots,$$

удовлетворяющей условиям 1) и 2), последовательности нижних ($s'_1, s'_2, \dots, s'_n, \dots$) и верхних ($S'_1, S'_2, \dots, S'_n, \dots$) интегральных сумм имеют тот же предел I .

Пусть

$$\lim_{n \rightarrow \infty} s'_n = \lim_{n \rightarrow \infty} S'_n = I'.$$

Рассмотрим последовательность разбиений отрезка $[a, b]$: $T''_1, T''_2, \dots, T''_n, \dots$, у которой каждое разбиение T''_n состоит из точек, входящих в разбиения T_n и T'_n . Эта последовательность так же удовлетворяет условиям 1) и 2).

Обозначая через s''_n и S''_n соответственно нижние и верхние интегральные суммы разбиения T''_n , на основании свойств интегральных сумм имеем:

$$s_n \leq s''_n \leq S''_n \leq S_n, \quad (4)$$

$$s'_n \leq s''_n \leq S''_n \leq S'_n. \quad (5)$$

Переходя в неравенствах (4) и (5) к пределу при $n \rightarrow \infty$, видим, что

$$\lim_{n \rightarrow \infty} s_n'' = \lim_{n \rightarrow \infty} S_n'' = I, \quad \lim_{n \rightarrow \infty} s_n' = \lim_{n \rightarrow \infty} S_n' = I',$$

т. е.

$$I = I'.$$

Это число I является пределом любых интегральных сумм при стремлении к нулю шага разбиения λ .

Действительно, пусть $\epsilon > 0$ — любое число. В силу равномерной непрерывности функции $f(x)$ на отрезке $[a, b]$ существует число $\delta > 0$, такое, что $|f(x') - f(x'')| < \frac{\epsilon}{b-a}$ для любых точек x' и x'' отрезка $[a, b]$, находящихся на расстоянии, меньшем δ .

Фиксируем какое-нибудь разбиение T с шагом, меньшим δ . Рассматривая его как начальное в последовательности разбиений, на основании неравенств (3)¹⁾ будем иметь:

$$s \leq I \leq S, \quad (6)$$

где

$$s = \sum_{i=1}^n m_i \Delta x_i, \quad S = \sum_{i=1}^n M_i \Delta x_i.$$

Любая интегральная сумма $\sigma = \sum_{i=1}^n f(\xi_i) \Delta x_i$ для того же разбиения заключена в пределах s и S [см. (1)]:

$$s \leq \sigma \leq S.$$

Сопоставляя неравенства (6) и (1), заключаем, что:

$$|\sigma - I| \leq S - s = \sum_{i=1}^n (M_i - m_i) \Delta x_i < \frac{\epsilon}{b-a} \sum_{i=1}^n \Delta x_i = \epsilon.$$

Так как T — любое, достаточно мелкое разбиение, то отсюда следует, что

$$\lim_{\lambda \rightarrow 0} \sigma = I,$$

т. е. функция $f(x)$ интегрируема на отрезке $[a, b]$.

Сформулируем в заключение более общую теорему существования определенного интеграла.

Теорема 2. Если функция $f(x)$ ограничена на отрезке $[a, b]$ и непрерывна на нем всюду, кроме конечного числа точек, то она интегрируема на этом отрезке.

Доказательство этой теоремы в данном курсе не рассматривается.

Возвращаясь к рассмотренным в § 1 задачам, видим, что:

1. Площадь S криволинейной трапеции, ограниченной отрезком $[a, b]$ оси абсцисс, прямыми $x=a$ и $x=b$ и кривой $y=f(x)$, где $f(x)$ — непрерывная, неотрицательная на отрезке $[a, b]$ функция, численно равна определенному интегралу от функции $f(x)$ на отрезке $[a, b]$:

$$S = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i = \int_a^b f(x) dx.$$

2. Работа направленной вдоль оси OX переменной по величине силы $F(|F|=f(x))$, где $f(x)$ — непрерывная функция) по перемещению единицы

¹⁾ Напомним, что $\underline{I} = \bar{I} = I$.

массы вдоль оси OX из точки $x=a$ в точку $x=b$ численно равна определенному интегралу от функции $f(x)$ на отрезке $[a, b]$:

$$A = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i = \int_a^b f(x) dx.$$

Первая задача позволяет установить *геометрический смысл определенного интеграла*. А именно, определенный интеграл от неотрицательной непрерывной функции $f(x)$ на отрезке $[a, b]$ численно равен площади криволинейной трапеции с основанием $[a, b]$, ограниченной сверху кривой $y=f(x)$ (рис. 243).

Заключение о механической интерпретации определенного интеграла, вытекающее из результата второй задачи, предоставляется сделать читателю.

§ 4. Дополнения к понятию определенного интеграла

При определении интеграла $\int_a^b f(x) dx$ предполагалось, что нижний предел интегрирования a меньше верхнего предела b . Распространим определение интеграла на случай, когда $a > b$. А именно, если $f(x)$ — функция, интегрируемая на отрезке $[b, a]$, то полагаем по определению, что

$$\int_a^b f(x) dx = - \int_b^a f(x) dx. \quad (1)$$

Примечание. Можно определить интеграл $\int_a^b f(x) dx$ при $a > b$ и как предел интегральной суммы при стремлении к нулю шага разбиения промежутка интегрирования от $x=a$ до $x=b$:

$$\int_a^b f(x) dx = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i.$$

При этом точки деления $x_0=a, x_1, x_2, \dots, x_n=b$ располагаются в направлении от a к b , т. е. $\Delta x_i = x_i - x_{i-1} < 0, i=1, 2, \dots, n$.

Если при построении интегральных сумм для интегралов $\int_a^b f(x) dx$ и $\int_b^a f(x) dx, a > b$, брать те же точки деления и те же точки ξ_i , то соответствующие интегральные суммы будут отличаться лишь знаком, а значит, в пределе получится равенство (1), которое мы взяли за определение интеграла при $a > b$.

В случае, когда $a=b$, для любой функции $f(x)$ полагаем, что

$$\int_a^a f(x) dx = 0.$$

§ 5. Свойства определенного интеграла

Установим, исходя из определения, некоторые простейшие свойства определенного интеграла. Будем при этом предполагать, что все рассматриваемые ниже функции на промежутках интегрирования удовлетворяют условиям теорем 1 или 2 и, следовательно, интегрируемы.

1. Значение определенного интеграла не зависит от обозначения переменной интегрирования, т. е.

$$\int_a^b f(x) dx = \int_a^b f(t) dt = \int_a^b f(u) du$$

и т. д.

Это утверждение непосредственно следует из определения интеграла как числа, равного пределу интегральных сумм. В случае неотрицательной подинтегральной функции утверждение 1 особенно очевидно, так как площадь соответствующей криволинейной трапеции не зависит от обозначения оси абсцисс.

2. Определенный интеграл меняет знак при перестановке пределов интегрирования, т. е. при $a \geq b$:

$$\int_a^b f(x) dx = - \int_b^a f(x) dx.$$

Доказательство следует из данного в § 4 дополнения к определению интеграла.

3.

$$\int_a^b dx = b - a \quad \text{при} \quad a \leq b.$$

Действительно, если $a < b$, то

$$\int_a^b dx = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \Delta x_i = \lim_{\lambda \rightarrow 0} (b - a) = b - a.$$

В случае $a \geq b$ по свойству 2 имеем:

$$\int_a^b dx = - \int_b^a dx = - (a - b) = b - a.$$

4. Постоянный множитель можно выносить за знак определенного интеграла, т. е.

$$\int_a^b kf(x) dx = k \int_a^b f(x) dx \quad \text{для любого числа } k.$$

Действительно, для любого разбиения T промежутка оси OX от $x=a$ до $x=b$ и любого выбора точек ξ_i

$$\sum_{i=1}^n kf(\xi_i) \Delta x_i = k \sum_{i=1}^n f(\xi_i) \Delta x_i.$$

Переходя к пределу при стремлении шага разбиения λ к нулю, получаем:

$$\begin{aligned} \int_a^b kf(x) dx &= \lim_{\lambda \rightarrow 0} \sum_{i=1}^n kf(\xi_i) \Delta x_i = \lim_{\lambda \rightarrow 0} k \sum_{i=1}^n f(\xi_i) \Delta x_i = \\ &= k \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i = k \int_a^b f(x) dx. \end{aligned}$$

5. Определенный интеграл от суммы двух функций $f_1(x)$ и $f_2(x)$ равен сумме определенных интегралов от слагаемых, т. е.

$$\int_a^b [f_1(x) + f_2(x)] dx = \int_a^b f_1(x) dx + \int_a^b f_2(x) dx.$$

Действительно, для любого разбиения промежутка от $x=a$ до $x=b$ и любого выбора точек ξ_i

$$\sum_{i=1}^n [f_1(\xi_i) + f_2(\xi_i)] \Delta x_i = \sum_{i=1}^n f_1(\xi_i) \Delta x_i + \sum_{i=1}^n f_2(\xi_i) \Delta x_i.$$

Так как

$$\lim_{\lambda \rightarrow 0} \sum_{i=1}^n f_1(\xi_i) \Delta x_i = \int_a^b f_1(x) dx,$$

$$\lim_{\lambda \rightarrow 0} \sum_{i=1}^n f_2(\xi_i) \Delta x_i = \int_a^b f_2(x) dx,$$

то получаем, что

$$\begin{aligned} \int_a^b [f_1(x) + f_2(x)] dx &= \lim_{\lambda \rightarrow 0} \sum_{i=1}^n [f_1(\xi_i) + f_2(\xi_i)] \Delta x_i = \\ &= \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f_1(\xi_i) \Delta x_i + \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f_2(\xi_i) \Delta x_i = \int_a^b f_1(x) dx + \int_a^b f_2(x) dx. \end{aligned}$$

Примечание. Свойство 5 имеет место для любого (фиксированного) числа слагаемых.

6. Если отрезок $[a, b]$ разбит точкой C на две части $[a, c]$ и $[c, b]$, то

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Доказательство. Так как интеграл $\int_a^b f(x) dx$, равный пределу интегральных сумм, не зависит от применяемого при составлении интегральных сумм способа разбиения отрезка $[a, b]$ на части, можно рассматривать только те разбиения, в которые точка c входит в качестве точки деления. Тогда

$$\sum_{i=1}^n f(\xi_i) \Delta x_i = \sum' f(\xi_i) \Delta x_i + \sum'' f(\xi_i) \Delta x_i,$$

где $\sum'$ и $\sum''$ — суммы, соответствующие отрезкам деления, попавшим соответственно на отрезки $[a, c]$ и $[c, b]$.

Переходя к пределу, при стремлении шага разбиения λ отрезка $[a, b]$ к нулю, получим:

$$\begin{aligned} \lambda_{[a, c]} \rightarrow 0, \quad \lambda_{[c, b]} \rightarrow 0, \\ \lim_{\lambda_{[a, c]} \rightarrow 0} \sum' f(\xi_i) \Delta x_i &= \int_a^c f(x) dx, \\ \lim_{\lambda_{[c, b]} \rightarrow 0} \sum'' f(\xi_i) \Delta x_i &= \int_c^b f(x) dx, \end{aligned}$$

а значит,

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

7. Для любого расположения точек a, b, c

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Для случая $a < c < b$ свойство 7 доказано выше.

Пусть, например, $a < b < c$. Тогда, применяя к отрезку $[a, c]$ свойство 6, получим:

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

По свойству 2

$$\int_b^c f(x) dx = -\int_c^b f(x) dx,$$

а значит,

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Аналогично доказывается свойство 7 для любого другого расположения точек a, b, c .

8. Изменение значений интегрируемой функции в конечном числе точек не влияет на величину интеграла.

Справедливость этого утверждения следует из интегрируемости данной и измененной функций, а также из того, что для любого сколь угодно мелкого разбиения отрезка интегрирования на части существуют интегральные суммы для этих функций, совпадающие между собой. Для построения таких сумм достаточно точки ξ_i выбирать так, чтобы они не совпадали с точками, в которых эти функции различны.

Примечание. Свойства 6—8 позволяют в дальнейшем отыскание предела интегральных сумм (короче — вычисление интеграла) для функции, имеющей конечное число точек разрыва первого рода на отрезке интегрирования, свести к вычислению суммы интегралов от непрерывных функций.

9. Если всюду на отрезке $[a, b]$ функция $f(x) \geq 0$, то

$$\int_a^b f(x) dx \geq 0.$$

В самом деле, любая интегральная сумма

$$\sum_{i=1}^n f(\xi_i) \Delta x_i$$

для функции $f(x)$ на отрезке $[a, b]$ неотрицательна, так как

$$f(\xi_i) \geq 0, \quad \Delta x_i = x_i - x_{i-1} > 0, \quad i = 1, 2, \dots, n.$$

Переходя к пределу в неравенстве

$$\sum_{i=1}^n f(\xi_i) \Delta x_i \geq 0,$$

получаем

$$\int_a^b f(x) dx \geq 0.$$

Геометрически утверждение очевидно.

10. Если всюду на отрезке $[a, b]$

$$f_1(x) \leq f_2(x),$$

то

$$\int_a^b f_1(x) dx \leq \int_a^b f_2(x) dx.$$

Действительно, так как на отрезке $[a, b]$ $f_2(x) - f_1(x) \geq 0$, то по свойству 9

$$\int_a^b [f_2(x) - f_1(x)] dx \geq 0.$$

По свойствам 5 и 4

$$\int_a^b [f_2(x) - f_1(x)] dx = \int_a^b f_2(x) dx - \int_a^b f_1(x) dx.$$

Следовательно,

$$\int_a^b f_2(x) dx - \int_a^b f_1(x) dx \geq 0, \quad \text{т. е.} \quad \int_a^b f_1(x) dx \leq \int_a^b f_2(x) dx.$$

Для неотрицательных на отрезке $[a, b]$ непрерывных функций $f_1(x)$ и $f_2(x)$ утверждение 9 очевидно, так как криволинейная трапеция с основанием $[a, b]$, соответствующая функции $f_2(x)$, содержит трапецию с тем же основанием, соответствующую функции $f_1(x)$.

11. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, неотрицательна и не равна тождественно нулю, то

$$\int_a^b f(x) dx > 0.$$

Доказательство. По условию найдется такая точка x_0 , $a \leq x_0 \leq b$, что $f(x_0) > 0$. В силу непрерывности функции $f(x)$ на отрезке $[a, b]$ существует отрезок $[\alpha, \beta]$, содержащий точку x_0 , на котором

$$f(x) > \frac{f(x_0)}{2} > 0.$$

Используя последовательно свойства 6, 9, 4, 3, получаем, что

$$\begin{aligned} \int_a^b f(x) dx &= \int_a^\alpha f(x) dx + \int_\alpha^\beta f(x) dx + \int_\beta^b f(x) dx \geq \int_\alpha^\beta f(x) dx \geq \\ &\geq \int_\alpha^\beta \frac{f(x_0)}{2} dx = \frac{f(x_0)}{2} (\beta - \alpha) > 0, \end{aligned}$$

т. е.

$$\int_a^b f(x) dx > 0.$$

Геометрически это утверждение очевидно.

Применяя свойство 11, можно доказать, что если функции $f_1(x)$ и $f_2(x)$ на отрезке $[a, b]$ непрерывны и не тождественно равны, то из неравенства

$$f_1(x) \leq f_2(x)$$

вытекает неравенство

$$\int_a^b f_1(x) dx < \int_a^b f_2(x) dx.$$

12. Для функции $f(x)$, заданной на отрезке $[a, b]$, имеет место неравенство

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

(модуль определенного интеграла от некоторой функции не превосходит интеграла от модуля этой функции).

Доказательство основано на использовании очевидных неравенств

$$-|f(x)| \leq f(x) \leq +|f(x)|,$$

справедливых для любого значения x отрезка $[a, b]$. Интегрируя их почленно (см. свойство 10) и применяя свойство 4, получаем неравенства

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq +\int_a^b |f(x)| dx,$$

откуда следует, что

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

Обобщенная теорема о среднем. Если функции $f(x)$ и $\varphi(x)$ непрерывны на отрезке $[a, b]$ и функция $\varphi(x)$ знакопостоянна на нем, то на этом отрезке существует точка ξ , такая, что

$$\int_a^b f(x) \varphi(x) dx = f(\xi) \int_a^b \varphi(x) dx.$$

Доказательство. Если $\varphi(x) \equiv 0$ на отрезке $[a, b]$, то утверждение теоремы очевидно. Пусть теперь $\varphi(x) \not\equiv 0$. Предположим, что на отрезке $[a, b]$ $\varphi(x) \geq 0$ (для случая $\varphi(x) \leq 0$ доказательство проводится аналогичным образом). Так как функция $f(x)$ непрерывна на отрезке $[a, b]$, то она принимает на нем свое наибольшее и наименьшее значения (обозначим их соответственно M и m), причем в любой точке отрезка $[a, b]$ $m \leq f(x) \leq M$.

Умножив на $\varphi(x)$ ($\varphi(x) \geq 0$!), получим неравенства

$$m\varphi(x) \leq f(x)\varphi(x) \leq M\varphi(x).$$

Проинтегрируем эти неравенства по отрезку $[a, b]$:

$$m \int_a^b \varphi(x) dx \leq \int_a^b f(x) \varphi(x) dx \leq M \int_a^b \varphi(x) dx. \quad (1)$$

Так как функция $\varphi(x)$ не тождественно равна нулю, неотрицательна и непрерывна на отрезке $[a, b]$, то

$$\int_a^b \varphi(x) dx > 0^1).$$

Разделим все части неравенств (1) на эту величину:

$$m \leq \frac{\int_a^b f(x) \varphi(x) dx}{\int_a^b \varphi(x) dx} \leq M.$$

Функция $f(x)$, непрерывная на отрезке $[a, b]$, между своими наименьшим и наибольшим значениями принимает все промежуточные, поэтому на отрезке $[a, b]$ найдется точка ξ ($a \leq \xi \leq b$), такая, что

$$\frac{\int_a^b f(x) \varphi(x) dx}{\int_a^b \varphi(x) dx} = f(\xi)$$

¹⁾ См. свойство 11.

или, что то же,

$$\int_a^b f(x) \varphi(x) dx = f(\xi) \int_a^b \varphi(x) dx^1).$$

Это и доказывает утверждение теоремы.

В частном случае, если $\varphi(x) \equiv 1$, из обобщенной теоремы о среднем вытекает, что

$$\int_a^b f(x) dx = f(\xi) \int_a^b dx, \quad a \leq \xi \leq b. \quad (2)$$

Так как

$$\int_a^b dx = b - a,$$

то равенство (2) запишется в виде:

$$\int_a^b f(x) dx = f(\xi) (b - a), \quad a \leq \xi \leq b.$$

Указанный частный случай также называют теоремой о среднем. Эта теорема о среднем может быть сформулирована следующим образом: определенный интеграл от функции $f(x)$, непрерывной на отрезке $[a, b]$, равен значению подинтегральной функции в некоторой «средней» точке ξ промежутка интегрирования, умноженному на длину промежутка интегрирования.

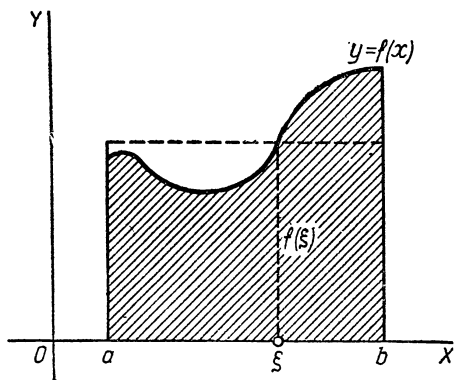


Рис. 247.

Если функция $f(x)$ неотрицательна на отрезке $[a, b]$, это утверждение означает, что площадь криволинейной трапеции с основанием $[a, b]$, соответствующей функции $f(x)$, численно равна площади прямоугольника с тем же основанием $[a, b]$ и высотой, равной значению функции $f(x)$ в некоторой «средней» точке ξ отрезка $[a, b]$ (см. рис. 247).

Примечание. Равенство

$$\int_a^b f(x) dx = f(\xi) (b - a),$$

очевидно, имеет место и в случае, когда $a > b$, т. е. для функции $f(x)$, непрерывной на отрезке $[b, a]$. При этом $b \leq \xi \leq a$.

§ 6. Определенный интеграл с переменным верхним пределом

Пусть $f(x)$ — функция, непрерывная на отрезке $[a, b]$. Тогда она непрерывна, а значит, интегрируема на любом отрезке $[a, x]$, где $a \leq x \leq b$.

Интеграл $\int_a^x f(t) dt^2)$ зависит от x , т. е. является функцией x , определенной на отрезке $[a, b]$:

$$\int_a^x f(t) dt = \Phi(x).$$

¹⁾ С помощью дополнительных рассуждений можно показать, что такую точку ξ можно найти внутри отрезка $[a, b]$.

²⁾ Переменная интегрирования обозначена буквой t , чтобы не было путаницы с переменным верхним пределом интеграла. Так как определенный интеграл от обозначения переменной интегрирования не зависит, можно тот же интеграл записать в виде:

$$\int_a^x f(x) dx.$$

Покажем, что функция $\Phi(x)$ дифференцируема на отрезке $[a, b]$. Для этого фиксируем любую точку x отрезка и дадим x приращение Δx ¹⁾. Функция $\Phi(x)$ получит приращение

$$\Delta\Phi(x) = \Phi(x + \Delta x) - \Phi(x) = \int_a^{x+\Delta x} f(t) dt - \int_a^x f(t) dt.$$

Так как

$$\int_a^{x+\Delta x} f(t) dt = \int_a^x f(t) dt + \int_x^{x+\Delta x} f(t) dt,$$

то

$$\Delta\Phi(x) = \int_x^{x+\Delta x} f(t) dt$$

или, по теореме о среднем,

$$\Delta\Phi(x) = f(\xi) \Delta x, \quad (1)$$

где ξ — точка, лежащая на отрезке с концами x и $x + \Delta x$.

Деля обе части равенства (1) на Δx и переходя к пределу при стремлении Δx к нулю, в силу непрерывности функции $f(x)$ получаем:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta\Phi(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} f(\xi) = f(x).$$

Следовательно, в точке x существует производная $\Phi'(x)$, причем $\Phi'(x) = f(x)$ ²⁾.

Таким образом, доказана теорема: производная определенного интеграла от непрерывной функции $f(x)$ по его верхнему пределу равна подинтегральной функции с заменой переменной интегрирования верхним пределом:

$$\frac{d}{dx} \int_a^x f(t) dt = f(x).$$

Аналогичным образом может быть доказана теорема о том, что производная определенного интеграла $\int_x^b f(t) dt$ от непрерывной на отрезке $[a, b]$ функции $f(x)$ по нижнему пределу равна взятой со знаком минус подинтегральной функции с заменой переменной интегрирования на нижний предел:

$$\frac{d}{dx} \int_x^b f(t) dt = -f(x).$$

Теорема о производной определенного интеграла по верхнему пределу позволяет доказать одну из основных теорем математического анализа — *теорему о существовании первообразной*.

Теорема. Для всякой функции $f(x)$, непрерывной на отрезке $[a, b]$, существует на этом отрезке первообразная, а значит, существует неопределенный интеграл.

Действительно, для непрерывной на отрезке $[a, b]$ функции $f(x)$, по теореме о существовании определенного интеграла, существует функция

$$\Phi(x) = \int_a^x f(t) dt, \quad a \leq x \leq b.$$

¹⁾ Приращение Δx берется так, чтобы точка $x + \Delta x$ лежала на отрезке $[a, b]$.

²⁾ При этом $\Phi'(a)$ — производная $\Phi(x)$ в точке a справа, а $\Phi'(b)$ — производная $\Phi(x)$ в точке b слева.

Так как $\Phi'(x) = f(x)$, то эта функция является первообразной для функции $f(x)$ на отрезке $[a, b]$.

Неопределенный интеграл от функции $f(x)$ на отрезке $[a, b]$ связан с определенным интегралом с переменным верхним пределом соотношением

$$\int f(x) dx = \int_a^x f(t) dt + C,$$

где C — произвольная постоянная.

§ 7. Формула Ньютона — Лейбница

Как было установлено выше, функция $f(x)$, непрерывная на отрезке $[a, b]$, имеет на этом отрезке первообразные, причем одной из них является функция

$$\Phi(x) = \int_a^x f(t) dt.$$

Пусть $F(x)$ — любая первообразная для функции $f(x)$ на том же отрезке $[a, b]$. Так как первообразные $\Phi(x)$ и $F(x)$ отличаются постоянным слагаемым, то имеет место равенство

$$\int_a^x f(t) dt = F(x) + C, \quad a \leq x \leq b,$$

где C — некоторое число.

Подставляя в это равенство значение $x = a$, будем иметь:

$$\int_a^a f(t) dt = F(a) + C, \quad 0 = F(a) + C, \quad C = -F(a).$$

Таким образом, для любого значения x , $a \leq x \leq b$,

$$\int_a^x f(t) dt = F(x) - F(a).$$

В частности, при $x = b$ получаем:

$$\int_a^b f(t) dt = F(b) - F(a).$$

Формула

$$\int_a^b f(x) dx = F(b) - F(a), \quad (1)$$

выражающая определенный интеграл от непрерывной функции через неопределенный, называется *формулой Ньютона — Лейбница*.

Разность $F(b) - F(a)$ принято условно записывать в виде:

$$F(x) \Big|_a^b \quad \text{или} \quad [F(x)]_a^b.$$

Формула (1) или, что то же, формула

$$\int_a^b f(x) dx = F(x) \Big|_a^b$$

показывает, что определенный интеграл от непрерывной функции равен разности значений любой ее первообразной, вычисленных для верхнего и нижнего пределов интегрирования.

Формула Ньютона — Лейбница дает простой и удобный метод вычисления определенных интегралов от непрерывных функций, применимый в тех случаях, когда первообразная подинтегральной функции может быть найдена в элементарных функциях.

Рассмотрим несколько примеров.

Пример 1. Вычислить интеграл $\int_0^2 (3x^2 - 1) dx$. Так как одной из первообразных для функции $f(x) = 3x^2 - 1$ является функция $x^3 - x$, то применяя формулу Ньютона — Лейбница, получаем:

$$\begin{aligned} \int_0^2 (3x^2 - 1) dx &= [x^3 - x]_0^2 = \\ &= (2^3 - 2) - (0^3 - 0) = 6. \end{aligned}$$

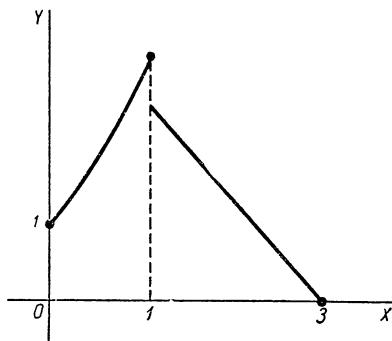


Рис. 248.

Если вместо первообразной $x^3 - x$ взять любую другую первообразную $x^3 - x + C$, где C — число, то результат, как этого и следует ожидать, получится прежним:

$$\int_0^2 (3x^2 - 1) dx = [x^3 - x + C]_0^2 = (2^3 - 2 + C) - (0^3 - 0 + C) = 6.$$

Пример 2.

$$\int_{-1}^1 \frac{dx}{x^2 + 1} = \operatorname{arctg} x \Big|_{-1}^1 = \operatorname{arctg} 1 - \operatorname{arctg} (-1) = \frac{\pi}{2}.$$

Пример 3. Функция

$$f(x) = \begin{cases} e^x, & 0 \leq x \leq 1, \\ 3 - x, & 1 < x \leq 3, \end{cases}$$

разрывна на отрезке $[0, 3]$ (рис. 248).

Используя замечание к свойству 8, можно свести вычисление интеграла $\int_0^3 f(x) dx$ к применению формулы Ньютона — Лейбница.

$$\begin{aligned} \int_0^3 f(x) dx &= \int_0^1 f(x) dx + \int_1^3 f(x) dx = \int_0^1 e^x dx + \int_1^3 (3 - x) dx = \\ &= e^x \Big|_0^1 + \left[3x - \frac{x^2}{2} \right]_1^3 = e - 1 + 9 - \frac{9}{2} - 3 + \frac{1}{2} = e + 1. \end{aligned}$$

Пример 4. Вычислим $\int_0^1 \sqrt{1 - x^2} dx$. Для этого находим первообразную функции $\sqrt{1 - x^2}$ на отрезке $[0, 1]$:

$$\begin{aligned} \int \sqrt{1 - x^2} dx &= \int \cos^2 t dt = \frac{1}{2} \int (1 + \cos 2t) dt = \frac{1}{2} t + \frac{1}{4} \sin 2t + C = \\ &= \frac{1}{2} \arcsin x + \frac{1}{2} x \sqrt{1 - x^2} + C^1). \end{aligned}$$

По формуле Ньютона — Лейбница

$$\int_0^1 \sqrt{1 - x^2} dx = \left[\frac{1}{2} \arcsin x + \frac{1}{2} x \sqrt{1 - x^2} \right]_0^1 = \frac{\pi}{4}.$$

¹⁾ Применялась подстановка $x = \sin t$.

В примере 4 для нахождения определенного интеграла нам пришлось сначала с помощью замены переменной вычислять неопределенный интеграл, затем возвращаться к переменной x и лишь затем применять формулу Ньютона—Лейбница. Вычисление этого и ряда других определенных интегралов можно упростить, если использовать формулу замены переменной в определенном интеграле, которая выводится в следующем параграфе.

§ 8. Замена переменной в определенном интеграле

Пусть требуется вычислить определенный интеграл $\int_a^b f(x) dx$ от функции $f(x)$, непрерывной на отрезке $[a, b]$. Имеет место следующее *правило замены переменной*.

Если: 1) функция $x = \varphi(t)$ непрерывна вместе со своей производной $\varphi'(t)$ на отрезке с концами $t = \alpha$, $t = \beta$ ($\alpha \leq \beta$) оси Ot ; 2) при изменении t от α до β значения функции $x = \varphi(t)$ не выходят за пределы отрезка $[a, b]$ и 3) $\varphi(\alpha) = a$, $\varphi(\beta) = b$, то

$$\int_a^b f(x) dx = \int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt. \quad (1)$$

Действительно, по формуле Ньютона—Лейбница

$$\int_a^b f(x) dx = F(b) - F(a),$$

где $F(x)$ — какая-нибудь первообразная для функции $f(x)$ на отрезке $[a, b]$.

Так как при этом функция $F(\varphi(t))$ является первообразной для функции $f(\varphi(t)) \varphi'(t)$, непрерывной на отрезке оси Ot с концами $t = \alpha$ и $t = \beta$, то по той же формуле Ньютона-Лейбница

$$\int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt = F(\varphi(\beta)) - F(\varphi(\alpha)).$$

Учитывая условия $\varphi(\beta) = b$, $\varphi(\alpha) = a$, заключаем, что

$$\int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt = F(b) - F(a),$$

а значит,

$$\int_a^b f(x) dx = \int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt.$$

Таким образом, если функция $x = \varphi(t)$ удовлетворяет указанным выше условиям 1, 2, 3, то подстановка $x = \varphi(t)$ сводит вычисление определенного интеграла

$$\int_a^b f(x) dx$$

к вычислению интеграла

$$\int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt.$$

Примечание. Вычисление интеграла $\int_a^b f(x) dx$ при $a > b$ сводится к рассмотренному случаю, так как

$$\int_a^b f(x) dx = - \int_b^a f(x) dx.$$

Пример 1. Вычислить $\int_0^1 \sqrt{1-x^2} dx$.

Рассмотрим подстановку $x = \sin t$. Здесь $\alpha = 0$, $\beta = \frac{\pi}{2}$, функция $x = \sin t$ и ее производная на отрезке $\left[0, \frac{\pi}{2}\right]$ непрерывны, при изменении t от $t=0$ до $t=\frac{\pi}{2}$ значения функции $x = \sin t$ не выходят за пределы отрезка $[0, 1]$. Применяя формулу (1), получаем:

$$\int_0^1 \sqrt{1-x^2} dx = \int_0^{\frac{\pi}{2}} \sqrt{1-\sin^2 t} \cos t dt = \int_0^{\frac{\pi}{2}} \cos^2 t dt = \frac{1}{2} \int_0^{\frac{\pi}{2}} (1 + \cos 2t) dt = \frac{\pi}{4}.$$

Формула (1), прочитанная справа налево, позволяет также сводить вычисление интеграла

$$\int_a^b f(\psi(x)) \psi'(x) dx$$

с помощью подстановки $t = \psi(x)$ к вычислению интеграла

$$\int_{\alpha}^{\beta} f(t) dt, \quad \text{где } \alpha = \psi(a), \quad \beta = \psi(b).$$

Пример 2. Вычислить $\int_0^1 \frac{e^x dx}{4e^{2x} + 12e^x + 34}$.

Применяя подстановку $t = e^x$, находим $\alpha = e^0 = 1$, $\beta = e^1 = e$,

$$\int_0^1 \frac{e^x dx}{4e^{2x} + 12e^x + 34} = \int_1^e \frac{dt}{4t^2 + 12t + 34} = \int_1^e \frac{dt}{(2t+3)^2 + 25}.$$

Для вычисления полученного интеграла можно применить подстановку $u = 2t+3$:

$$\int_1^e \frac{dt}{(2t+3)^2 + 25} = \frac{1}{2} \int_5^{2e+3} \frac{du}{u^2 + 25} = \frac{1}{10} \operatorname{arctg} \frac{u}{5} \Big|_5^{2e+3} = \frac{1}{10} \operatorname{arctg} \frac{2e+3}{5} - \frac{\pi}{40}.$$

Таким образом,

$$\int_0^1 \frac{e^x dx}{4e^{2x} + 12e^x + 34} = \frac{1}{10} \operatorname{arctg} \frac{2e+3}{5} - \frac{\pi}{40}.$$

Примечание. При использовании формулы замены переменной в определенном интеграле необходимо проверять выполнение перечисленных выше условий.

Если эти условия нарушаются, замена переменной по указанной формуле может привести к абсурду.

Пример 3. $\int_0^{\pi} dx = x \Big|_0^{\pi} = \pi$.

С другой стороны,

$$\int_0^{\pi} dx = \int_0^{\pi} \frac{dx}{\sin^2 x + \cos^2 x} = \int_0^{\pi} \frac{dx}{\cos^2 x (1 + \operatorname{tg}^2 x)}.$$

Подстановка $\operatorname{tg} x = t$ формально приводит к следующему результату:

$$\int_0^{\pi} dx = \int_0^{\pi} \frac{dx}{\cos^2 x (1 + \operatorname{tg}^2 x)} = \int_0^0 \frac{dt}{1+t^2} = 0, \quad \text{т. е. } \pi = 0.$$

В данном случае подстановка $\operatorname{tg} x = t$ неприменима.

§ 9. Формула интегрирования по частям в определенном интеграле

Пусть функции $u = \varphi(x)$ и $v = \psi(x)$ непрерывны вместе со своими производными на отрезке $[a, b]$ ¹⁾.

Проинтегрируем тождество

$$[\varphi(x) \cdot \psi(x)]' = \varphi'(x) \psi(x) + \varphi(x) \psi'(x)$$

в пределах от $x=a$ до $x=b$:

$$\int_a^b [\varphi(x) \cdot \psi(x)]' dx = \int_a^b \varphi'(x) \psi(x) dx + \int_a^b \varphi(x) \psi'(x) dx. \quad (1)$$

Так как функция $\varphi(x) \cdot \psi(x)$ является первообразной для функции $[\varphi(x) \cdot \psi(x)]'$, то по формуле Ньютона — Лейбница:

$$\int_a^b [\varphi(x) \cdot \psi(x)]' dx = \varphi(x) \cdot \psi(x) \Big|_a^b.$$

Равенство (1) при этом запишется в виде:

$$\varphi(x) \cdot \psi(x) \Big|_a^b = \int_a^b \varphi'(x) \psi(x) dx + \int_a^b \varphi(x) \psi'(x) dx$$

или, что то же, в виде:

$$\int_a^b \varphi(x) \psi'(x) dx = \varphi(x) \cdot \psi(x) \Big|_a^b - \int_a^b \varphi'(x) \psi(x) dx. \quad (2)$$

Соотношение (2) называется *формулой интегрирования по частям* в определенном интеграле.

Эту формулу обычно записывают более компактно следующим образом:

$$\int_a^b u dv = uv \Big|_a^b - \int_a^b v du.$$

Пример. Вычислить

$$\int_0^{\frac{\pi}{2}} x \sin x dx.$$

Обозначая $x=u$, $\sin x dx = dv$ (тогда $dx=du$, $-\cos x=v$), по формуле интегрирования по частям получим:

$$\int_0^{\frac{\pi}{2}} x \sin x dx = -x \cos x \Big|_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos x dx = \int_0^{\frac{\pi}{2}} \cos x dx = \sin x \Big|_0^{\frac{\pi}{2}} = 1.$$

Раздел 2

ГЕОМЕТРИЧЕСКИЕ ПРИЛОЖЕНИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

§ 10. Площадь плоской фигуры

1. Общие сведения. В элементарной геометрии дается определение площади треугольника, а также площади любого многоугольника как суммы площадей составляющих его треугольников.

Дадим определение площади любой плоской фигуры так, чтобы оно как частный случай включало в себя эти известные определения.

¹⁾ Или на отрезке $[b, a]$.

Пусть σ — любая фигура, лежащая в плоскости XOY .

Всякий многоугольник Q , содержащий фигуру σ , будем называть *описанным* около фигуры σ , а всякий многоугольник q , содержащийся в фигуре σ , будем называть *вписанным* в фигуру σ (рис. 249).

Если существуют последовательности описанных и вписанных в фигуру σ многоугольников (обозначим их соответственно $Q_1, Q_2, \dots, Q_n \dots$ и $q_1, q_2, \dots, q_n, \dots$), такие, что последовательности их площадей имеют общий предел (конечный)

$$\lim_{n \rightarrow \infty} \text{пл. } Q_n = \lim_{n \rightarrow \infty} \text{пл. } q_n = S,$$

то фигура σ называется *квадрируемой*, а этот общий предел площадей описанных и вписанных многоугольников называется *площадью* фигуры:

$$\text{пл. } \sigma = S^1).$$

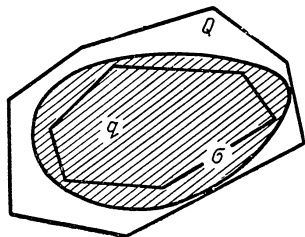


Рис. 249.

Из этого определения, в частности, следует, что если существуют последовательности описанных и вписанных в фигуру σ квадрируемых фигур (необязательно многоугольников), площади которых имеют общий (конечный) предел S , то фигура квадрируема и ее площадь равна этому числу S .

Можно также показать, что если фигуру σ разбить на две части σ_1 и σ_2 (рис. 250), то из квадрируемости любых двух из этих фигур следует квадрируемость третьей, причем всегда

$$\text{пл. } \sigma = \text{пл. } \sigma_1 + \text{пл. } \sigma_2.$$

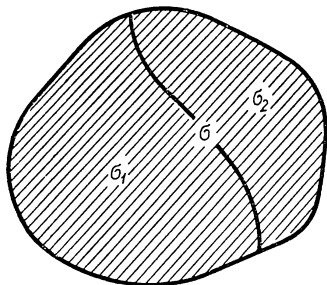


Рис. 250.

2. Площадь криволинейной трапеции.

Выше мы решали задачу о вычислении площади криволинейной трапеции, ограниченной снизу отрезком $[a, b]$ оси OX , сверху — кривой $y = f(x)$, где $f(x)$ — непрерывная, неотрицательная на отрезке $[a, b]$ функция, слева и справа — прямыми $x = a$ и $x = b$. Покажем, что такая криволинейная трапеция квадрируема и что данное ранее определение площади криволинейной трапеции совпадает с общим определением площади.

Рассмотрим любое разбиение T отрезка $[a, b]$ на части $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ (здесь $a = x_0, b = x_n$). Обозначая через m_i и M_i соответственно наименьшее и наибольшее значения функции $f(x)$ на отрезке $[x_{i-1}, x_i]$, составим для данного разбиения T верхнюю и нижнюю интегральные суммы:

$$S = \sum_{i=1}^n M_i \Delta x_i, \quad s = \sum_{i=1}^n m_i \Delta x_i.$$

Эти суммы равны площадям ступенчатых фигур, соответственно описанных и вписанных в данную криволинейную трапецию (рис. 245).

По теореме существования определенного интеграла при стремлении к нулю шага разбиения отрезка $[a, b]$ получаем:

$$\lim_{\lambda \rightarrow 0} \sum_{i=1}^n M_i \Delta x_i = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n m_i \Delta x_i = \int_a^b f(x) dx.$$

¹⁾ У любой другой последовательности описанных и вписанных в такую фигуру многоугольников, площади которых имеют общий предел, этот предел будет тем же.

Отсюда следует, что для любой последовательности разбиений отрезка $[a, b]$ с шагом разбиения, стремящимся к нулю, площади описанных и вписанных ступенчатых фигур (равные соответственно верхним и нижним интегральным суммам) стремятся к общему пределу, равному определенному интегралу:

$$\int_a^b f(x) dx.$$

Таким образом, данная криволинейная трапеция квадратуема и ее площадь S вычисляется по формуле

$$S = \int_a^b f(x) dx \quad (1)$$

или короче:

$$S = \int_a^b y dx.$$

Если функция $f(x)$ непрерывна и неположительная на отрезке $[a, b]$, то $\int_a^b f(x) dx \leq 0$.

Абсолютная величина этого интеграла равна площади криволинейной трапеции, ограниченной сверху отрезком $[a, b]$ оси OX , слева и справа — прямыми $x=a$ и $x=b$, снизу — кривой $y=f(x)$ (рис. 251).

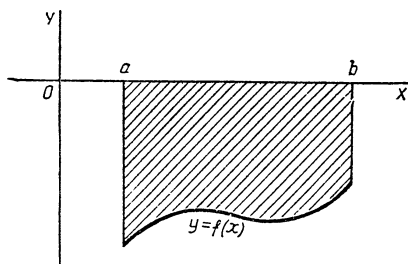


Рис. 251.

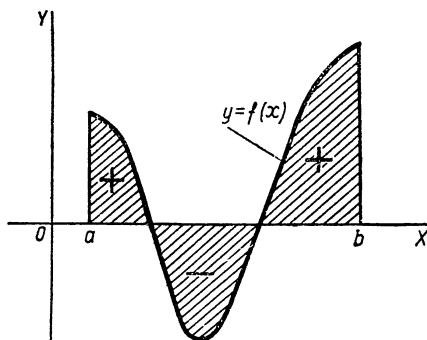


Рис. 252.

Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и не знакопостоянная, то определенный интеграл $\int_a^b f(x) dx$ численно равен алгебраической сумме площадей криволинейных трапеций, лежащих над и под осью OX ¹⁾. В эту сумму площади криволинейных трапеций, лежащих над осью OX , входят со знаком плюс, а площади криволинейных трапеций, лежащих под осью OX , со знаком минус (рис. 252).

Площадь фигуры $ABCD$, ограниченной слева и справа прямыми $x=a$ и $x=b$, снизу и сверху — кривыми $y=\varphi(x)$ и $y=\psi(x)$, где $\varphi(x)$ и $\psi(x)$ — непрерывные функции на отрезке $[a, b]$, можно вычислить как разность площадей криволинейных трапеций $LKCD$ и $LKBA$ (рис. 253):

$$S = \int_a^b \psi(x) dx - \int_a^b \varphi(x) dx,$$

т. е.

$$S = \int_a^b [\psi(x) - \varphi(x)] dx. \quad (2)$$

¹⁾ Предполагается, что число перемен знака у функции $f(x)$ на отрезке $[a, b]$ конечно.

Площади квадратуемых фигур более сложного вида обычно вычисляются с помощью определенного интеграла после предварительного разбиения их на части указанного выше вида. При этом предполагается, что рассматриваемую фигуру можно представить в виде алгебраической суммы конечного числа криволинейных трапеций.

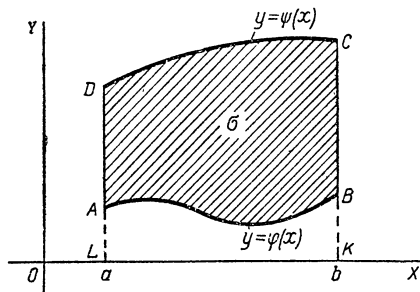


Рис. 253.

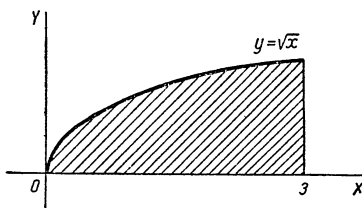


Рис. 254.

Приведем несколько примеров на вычисление площадей плоских фигур.

Пример 1. Вычислить площадь фигуры, ограниченной линиями $y = \sqrt{x}$, $x = 3$, $y = 0$ (рис. 254).

Так как данная фигура представляет собой криволинейную трапецию, лежащую над осью OX , то

$$S = \int_0^3 \sqrt{x} \, dx = \frac{2}{3} x \sqrt{x} \Big|_0^3 = 2\sqrt{3}.$$

Пример 2. Найти площадь фигуры, ограниченной линиями $y = x$, $y = x^2$, $x = 2$ (рис. 255).

Применяя формулу (2), получаем:

$$S = \int_1^2 (x^2 - x) \, dx = \left[\frac{x^3}{3} - \frac{x^2}{2} \right]_1^2 = \frac{5}{6}.$$

Пример 3. Найти площадь фигуры, ограниченной осью OX и кривой $y = \cos x$, $0 \leq x \leq \frac{3\pi}{2}$ (рис. 256).

Так как функция $\cos x$ на отрезке $[0, \frac{\pi}{2}]$ неотрицательна, а на отрезке $[\frac{\pi}{2}, \frac{3\pi}{2}]$ неположительна, получаем:

$$S = \int_0^{\frac{\pi}{2}} \cos x \, dx - \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \cos x \, dx = \sin x \Big|_0^{\frac{\pi}{2}} - \sin x \Big|_{\frac{\pi}{2}}^{\frac{3\pi}{2}} = 3.$$

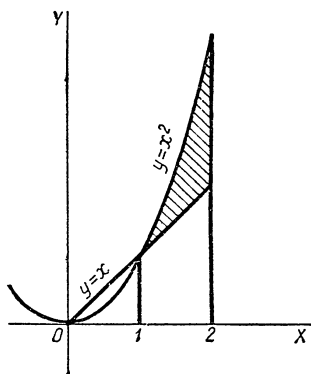


Рис. 255.

Пример 4. Найти площадь фигуры, ограниченной осью OX и одной аркой циклоиды: $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $0 \leq t \leq 2\pi$ (рис. 257).

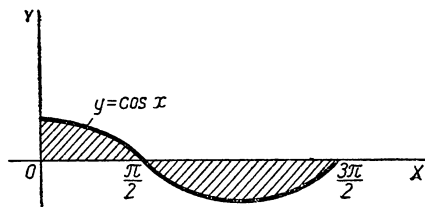


Рис. 256.

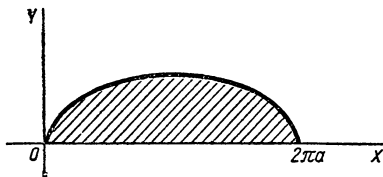


Рис. 257.

Применяя формулу (1), получим: $S = \int_0^{2\pi a} y \, dx$.

Произведем в полученном интеграле замену переменной по формуле $x = a(t - \sin t)$. Тогда (по условию) $y = a(1 - \cos t)$. При $x = 0$ имеем $t = 0$, при $x = 2\pi a$ имеем

$t = 2\pi$. Таким образом,

$$S = \int_0^{2\pi} y dx = \int_0^{2\pi} a^2 (1 - \cos t)^2 dt = a^2 \int_0^{2\pi} (1 - 2 \cos t + \cos^2 t) dt = \\ = a^2 \left[t - 2 \sin t + \frac{t}{2} + \frac{\sin 2t}{4} \right]_0^{2\pi} = 3\pi a^2.$$

3. Площадь криволинейного сектора. Выведем теперь формулу для вычисления площади криволинейного сектора.

Введем на плоскости полярную систему координат и рассмотрим криволинейный сектор $OABO$, ограниченный двумя лучами OA и OB , составляющими с полярной осью углы α и β ($\alpha < \beta$), и кривой AB , уравнение которой в полярных координатах (с полюсом O) имеет вид: $\rho = f(\varphi)$, где $f(\varphi)$ — неотрицательная непрерывная функция для $\alpha \leq \varphi \leq \beta$ (рис. 258).

Определение площади такой фигуры можно произвести следующим образом.

Разобьем угол AOB лучами $\varphi = \varphi_0, \varphi = \varphi_1, \dots, \varphi = \varphi_n$, где $\varphi_0 = \alpha < \varphi_1 < \varphi_2 < \dots < \varphi_n = \beta$, на n частей.

Обозначим точки пересечения этих лучей с кривой $\rho = f(\varphi)$ соответственно через $A_0, A_1, \dots, A_n$ (A_0 совпадает с A , а A_n с B).

При этом криволинейный сектор $OABO$ разобьется на n криволинейных секторов. Углы $\widehat{A_0OA_1}, \widehat{A_1OA_2}, \dots, \widehat{A_{n-1}OA_n}$ соответственно будут равны:

$$\Delta\varphi_1 = \varphi_1 - \varphi_0, \quad \Delta\varphi_2 = \varphi_2 - \varphi_1, \quad \dots, \quad \Delta\varphi_n = \varphi_n - \varphi_{n-1}.$$

Для каждого сектора $OA_{i-1}A_iO, i = 1, 2, \dots, n$, определив наибольшее и наименьшее значения функции $\rho = f(\varphi)$ на промежутке $[\varphi_{i-1}, \varphi_i]$

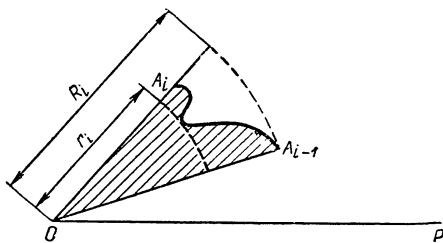


Рис. 259.

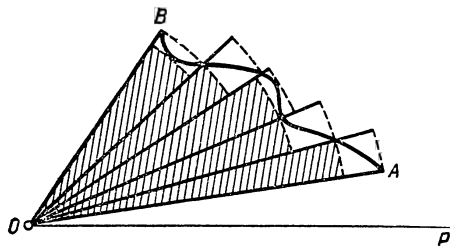


Рис. 260.

(обозначим их R_i и r_i), построим круговые секторы радиусов R_i и r_i с тем же углом при вершине (рис. 259). Площади этих круговых секторов соответственно равны:

$$\frac{R_i^2}{2} \Delta\varphi_i \quad \text{и} \quad \frac{r_i^2}{2} \Delta\varphi_i.$$

Фигуру Q , составленную из всех круговых секторов радиусов $R_i, i = 1, 2, \dots, n$, назовем описанной вокруг данного криволинейного сектора, она содержит этот сектор. Фигуру q , составленную из всех круговых секторов радиусов $r_i, i = 1, 2, \dots, n$, назовем вписанной. Эта фигура содержится в данном криволинейном секторе (рис. 260).

При этом

$$\text{пл. } Q = \sum_{i=1}^n \frac{R_i^2}{2} \Delta\varphi_i, \quad \text{пл. } q = \sum_{i=1}^n \frac{r_i^2}{2} \Delta\varphi_i.$$

Так как полученные суммы представляют собой интегральные суммы (верхнюю и нижнюю) для непрерывной на промежутке $[\alpha, \beta]$ функции $\frac{f^2(\varphi)}{2}$, то при стремлении наибольшего из углов деления (назовем его шагом разбиения указанного промежутка и обозначим через λ) к нулю эти суммы стремятся к общему пределу:

$$\lim_{\lambda \rightarrow 0} \sum_{i=1}^n \frac{R_i^2}{2} \Delta \varphi_i = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \frac{r_i^2}{2} \Delta \varphi_i = \frac{1}{2} \int_{\alpha}^{\beta} f^2(\varphi) d\varphi.$$

Отсюда следует, что для любой последовательности разбиений угла AOB на части с шагом разбиения λ , стремящимся к нулю, площади соответствующих описанных и вписанных фигур имеют общий предел,

равный $\frac{1}{2} \int_{\alpha}^{\beta} f^2(\varphi) d\varphi$, т. е. следует, что дан-

ный криволинейный сектор есть квадратуемая фигура и что его площадь можно вычислять по формуле:

$$S = \frac{1}{2} \int_{\alpha}^{\beta} f^2(\varphi) d\varphi$$

или короче:

$$S = \frac{1}{2} \int_{\alpha}^{\beta} \rho^2 d\varphi. \quad (3)$$

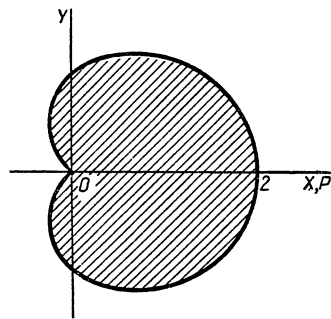


Рис. 261.

Пример. Найти площадь фигуры, ограниченной кардиоидой $\rho = 1 + \cos \varphi$ (рис. 261).

Применяя формулу (3), получаем

$$S = \frac{1}{2} \int_0^{2\pi} (1 + \cos \varphi)^2 d\varphi = \frac{3\pi}{2}.$$

§ 11. Объем тела

1. Общие сведения. В элементарной геометрии дано определение объема любого многогранника как суммы объемов составляющих его пирамид. Если тело ограничено произвольной поверхностью, то его, вообще говоря, нельзя составить из конечного числа пирамид. Обобщим понятие объема на тела произвольной формы.

Рассмотрим в пространстве тело V . Всякий многогранник K , содержащий внутри это тело, назовем *описанным*, а всякий многогранник k , содержащийся внутри тела V , назовем *вписанным* в это тело. Будем говорить, что тело V *кубируемо*, если существуют последовательности описанных и вписанных в него многогранников: $K_1, K_2, \dots, K_n, \dots$ и $k_1, k_2, \dots, k_n, \dots$, объемы которых имеют общий (конечный) предел:

$$\lim_{n \rightarrow \infty} \text{об. } K_n = \lim_{n \rightarrow \infty} \text{об. } k_n.$$

Этот общий предел объемов описанных и вписанных в тело V многогранников называется *объемом* тела V .

Можно показать, что если тело V разбито на две части V_1 и V_2 , то из кубируемости любых двух из этих тел следует кубируемость третьего, причем $\text{об. } V = \text{об. } V_1 + \text{об. } V_2$.

Из данного выше определения следует, что если для тела V существуют две последовательности описанных и вписанных в него кубируемых тел (не обязательно многогранников), объемы которых имеют общий

(конечный) предел, то тело V также кубируемо и его объем равен указанному общему пределу объемов описанных и вписанных тел.

Примером кубируемого тела является цилиндр с любым квадратуемым основанием σ (рис. 262).

Действительно, так как основание цилиндра квадратуемо, существуют последовательности многоугольников, описанных и вписанных в фигуру σ , обозначим их

$$Q_1, Q_2, \dots, Q_n, \dots \text{ и } q_1, q_2, \dots, q_n, \dots,$$

такие, что

$$\lim_{n \rightarrow \infty} \text{пл. } Q_n = \lim_{n \rightarrow \infty} \text{пл. } q_n = S.$$

Рассмотрим последовательности многогранников

$$K_1, K_2, \dots, K_n, \dots \text{ и } k_1, k_2, \dots, k_n, \dots,$$

основаниями которых являются указанные выше описанные и вписанные в σ многоугольники, боковые грани параллельны образующей цилиндра, а высоты равны высоте H цилиндра.

Многогранники K_n и k_n будут соответственно описанными и вписанными в данное тело. Их объемы находим по формулам

$$\text{об. } K_n = H \cdot \text{пл. } Q_n, \quad \text{об. } k_n = H \cdot \text{пл. } q_n.$$

Так как

$$\lim_{n \rightarrow \infty} H \cdot \text{пл. } Q_n = \lim_{n \rightarrow \infty} H \cdot \text{пл. } q_n = H \cdot S,$$

то цилиндр с основанием σ и высотой H кубируем и его объем равен произведению площади основания на высоту.

2. Вычисление объема тела по известным площадям его поперечных сечений. Рассмотрим одну частную задачу на вычисление объема тела. Общая задача об объеме тела решается при изучении кратных интегралов (см. разделы 3 и 4, гл. IX).

Пусть дано тело V , у которого площадь любого сечения σ , перпендикулярного некоторой оси (примем ее за ось OX), известна как непрерывная функция x :

$$\text{пл. } \sigma = S(x), \quad a \leq x \leq b.$$

Абсциссы $x=a$ и $x=b$ соответствуют крайним точкам тела V (рис. 263).

Предположим, что тело V таково, что любая пара его сечений плоскостями, перпендикулярными оси OX , проектируется одно в другое на некоторую плоскость, перпендикулярную оси OX .

Покажем, что V тело кубируемо и что его объем ¹⁾ можно вычислить по формуле: $V = \int_a^b S(x) dx$.

Для решения поставленной задачи разобьем отрезок $[a, b]$ точками $x_0=a, x_1, x_2, \dots, x_n=b$ ($x_0 < x_1 < \dots < x_n$) на n частей и проведем через точки деления плоскости, перпендикулярные оси OX . При этом тело V разобьется на n «дисков».

Так как на отрезке $[x_{i-1}, x_i]$ функция $S(x)$ непрерывна, то она принимает на нем наименьшее и наибольшее значения. Обозначим эти значения через m_i и M_i .

Если все сечения i -го «диска» спроектировать на плоскость $x=x_i$, то они будут содержать в себе сечение с площадью m_i и содержаться в сечении с площадью M_i .

¹⁾ Условимся объем тела V обозначать той же буквой, что и само тело.

Построим на этих наименьшем и наибольшем сечениях цилиндры с образующими, параллельными оси OX , высоты $\Delta x_i = x_i - x_{i-1}$ (рис. 264). «Большой» из них будет описан около i -го «диска», а «меньший» — вписан в него. Объемы описанного и вписанного цилиндров будут соответственно равны:

$$M_i \Delta x_i \quad \text{и} \quad m_i \Delta x_i.$$

Производя указанные построения над каждым из отрезков деления, получим описанное около тела V и вписанное в него тела, составленные из n цилиндров. Их объемы равны:

$$\sum_{i=1}^n M_i \Delta x_i \quad \text{и} \quad \sum_{i=1}^n m_i \Delta x_i.$$

При стремлении к нулю шага разбиения λ отрезка $[a, b]$ эти суммы имеют общий предел:

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \sum_{i=1}^n M_i \Delta x_i &= \lim_{\lambda \rightarrow 0} \sum_{i=1}^n m_i \Delta x_i = \\ &= \int_a^b S(x) dx. \end{aligned}$$

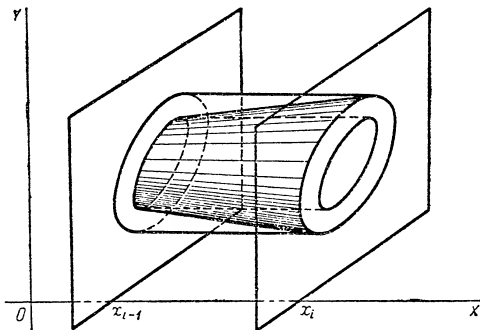


Рис. 264.

Рассматривая последовательность разбиений отрезка $[a, b]$ на части с шагом разбиения, стремящимся к нулю, заключаем, что существуют последовательности описанных и вписанных в тело кубирuemых тел, объемы которых имеют общий предел.

Отсюда следует, что тело V кубируемо и его объем можно вычислять по формуле

$$V = \int_a^b S(x) dx. \quad (1)$$

3. Объем тела вращения. Рассмотрим на плоскости XOY кривую $y = f(x)$, где $f(x)$ — функция, непрерывная на отрезке $[a, b]$. Если соответствующую ей криволинейную трапецию с основанием $[a, b]$ вращать вокруг оси OX , то образуется так называемое *тело вращения* (рис. 265). Это тело удовлетворяет всем высказанным при выводе формулы (3) § 10 условиям, причем

$$S(x) = \pi f^2(x),$$

так как каждое сечение тела плоскостью $x = \text{const}$, перпендикулярной оси OX , есть круг радиуса $R = |f(x)|$.

Применяя формулу (1), найдем объем тела вращения:

$$V = \pi \int_a^b f^2(x) dx$$

или короче:

$$V = \pi \int_a^b y^2 dx. \quad (2)$$

Пример 1. Найти объем тела, ограниченного поверхностью, полученной вращением кривой $y = x^3$, $0 \leq x \leq 1$, вокруг оси OX (рис. 266).

На основании формулы (2)

$$V = \pi \int_0^1 x^6 dx = \frac{\pi}{7}.$$

Пример 2. Найти объем тела, ограниченного поверхностью, полученной в результате вращения вокруг оси OX одной арки циклоиды: $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $0 \leq t \leq 2\pi$ (рис. 267).

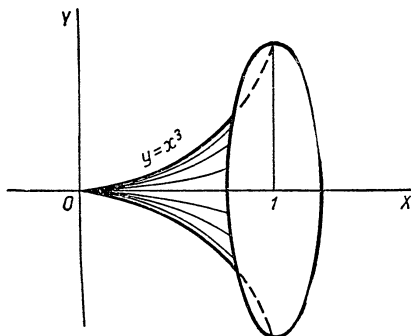


Рис. 266.

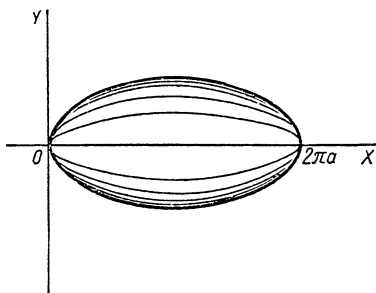


Рис. 267.

По формуле (2)

$$V = \pi \int_0^{2\pi a} y^2 dx.$$

Произведя в этом интеграле замену переменных $x = a(t - \sin t)$, получим:

$$\begin{aligned} V &= \pi a^3 \int_0^{2\pi} (1 - \cos t)^2 dt = \pi a^3 \int_0^{2\pi} \left[1 - 3 \cos t + 3 \frac{1 + \cos 2t}{2} + \cos t (1 - \sin^2 t) \right] dt = \\ &= \pi a^3 \left[t - 3 \sin t + \frac{3}{2} t + \frac{3}{4} \sin 2t + \sin t - \frac{\sin^3 t}{3} \right]_0^{2\pi} = 5\pi^2 a^3. \end{aligned}$$

§ 12. Длина дуги кривой

1. Общие понятия. В элементарной геометрии вводятся понятия длины отрезка, длины ломаной линии (как суммы длин ее звеньев), длины окружности (как предела периметров правильных вписанных в нее многоугольников при неограниченном увеличении числа их сторон) и ее частей. Основываясь на понятии длины отрезка, дадим определение длины произвольной плоской и пространственной кривой таким образом, чтобы оно включало в себя как частный случай уже известные определения.

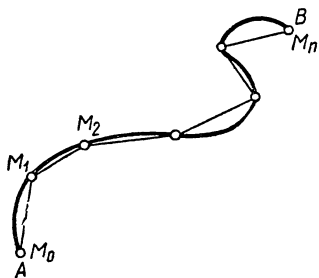


Рис. 268.

Пусть AB — несамопересекающаяся незамкнутая кривая на плоскости или в пространстве. Разобьем ее точками $M_1, M_2, \dots, M_{n-1}$, занумерованными в направлении от A к B , на n дуг. Обозначим для удобства записи точку A через M_0 , а точку B через

M_n . Соединив точки деления, получим ломаную линию $M_0 M_1 \dots M_n$, вписанную в дугу AB (рис. 268). Ее длина p равна сумме длин всех звеньев. Чем меньше длины звеньев, тем меньше ломаная $M_0 M_1 \dots M_n$ по своей форме отличается от кривой AB . За длину s кривой AB принимается предел, к которому стремятся периметры вписанных в эту кривую ломаных при стремлении к нулю наибольших из длин звеньев этих ломаных [если такой предел (конечный) существует]. При этом число s

называют *пределом* периметров ломаных, вписанных в кривую AB , при указанном выше процессе, если для любого числа $\varepsilon > 0$ существует число $\delta > 0$, такое, что для любой вписанной в кривую AB ломаной с длинами звеньев, меньшими δ , имеет место неравенство

$$|s - p| < \varepsilon.$$

Отсюда следует, что число s не зависит от выбора вписанных в кривую ломаных, длины звеньев которых стремятся к нулю.

Те кривые, для которых этот предел s существует, называются *спрямляемыми*.

Ниже будут приведены достаточные условия спрямляемости кривых и выведены формулы для вычисления длин кривых.

Из данных выше определений следует, что если кривая AB точкой C разбита на две части AC и CB , то из спрямляемости любых двух из дуг AC , CB и AB следует спрямляемость третьей, причем всегда

$$\text{дл. } AB = \text{дл. } AC + \text{дл. } CB.$$

Если кривая AB замкнута или самопересекается, данное выше определение длины дуги не имеет смысла, так как в этих случаях из стремления к нулю длин всех звеньев вписанных в кривую ломаных не следует, что эти ломаные по своей форме приближаются к данной кривой (см. рис. 269).

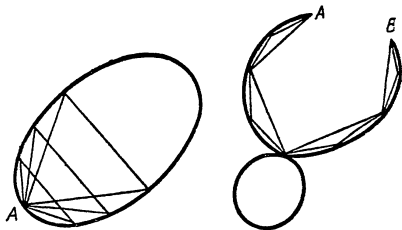


Рис. 269.

Замкнутую, а также самопересекающуюся кривую AB условимся называть *спрямляемой*, если ее можно разбить на конечное число спрямляемых дуг рассмотренного выше вида. *Длину* такой кривой полагаем равной сумме длин составляющих ее «простых» дуг.

2. Длина дуги плоской кривой. Рассмотрим в плоскости XOY кривую AB^1), заданную уравнениями в параметрической форме:

$$x = \varphi(t), \quad y = \psi(t), \quad \alpha \leq t \leq \beta. \quad (1)$$

Будем предполагать, что изменению параметра t от α к β соответствует движение точки (x, y) , $x = \varphi(t)$, $y = \psi(t)$, по кривой в направлении от A к B . Если при этом функции $\varphi(t)$ и $\psi(t)$ непрерывны вместе со своими производными на отрезке $[\alpha, \beta]$, то кривая AB спрямляема и ее длина s вычисляется по формуле

$$s = \int_{\alpha}^{\beta} \sqrt{\varphi'^2(t) + \psi'^2(t)} dt.$$

Для доказательства высказанного утверждения разобьем дугу AB точками $M_1, \dots, M_{n-1}$, занумерованными в направлении от A к B , на части. Обозначим точку A через M_0 , а точку B через M_n . Каждой точке $M_i(x_i, y_i)$ соответствует определенное значение параметра t_i : $x_i = \varphi(t_i)$, $y_i = \psi(t_i)$. Точки $t_0 = \alpha$, $t_1, \dots, t_n = \beta$ разобьют отрезок $[\alpha, \beta]$ оси Ot на n частей. Впишем в дугу AB ломаную $M_0 M_1 \dots M_n$. Длина p этой ломаной равна сумме длин составляющих ее звеньев:

$$\begin{aligned} p &= \sum_{i=1}^n M_{i-1} M_i = \sum_{i=1}^n \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} = \\ &= \sum_{i=1}^n \sqrt{[\varphi(t_i) - \varphi(t_{i-1})]^2 + [\psi(t_i) - \psi(t_{i-1})]^2}. \end{aligned}$$

¹⁾ Кривая AB предполагается незамкнутой и несамопересекающейся.

Применяя к разности в каждой квадратной скобке теорему Лагранжа, находим:

$$p = \sum_{i=1}^n \sqrt{\varphi'^2(\tau_i) + \psi'^2(\tau_i^*)} \Delta t_i,$$

где

$$\Delta t_i = t_i - t_{i-1}, \quad t_{i-1} < \tau_i < t_i, \quad t_{i-1} < \tau_i^* < t_i.$$

Полученная сумма не является интегральной (значения стоящих под корнем функций вычисляются в двух, вообще говоря, различных точках каждого отрезка деления).

Рассмотрим сумму $\tilde{p} = \sum_{i=1}^n \sqrt{\varphi'^2(\tau_i) + \psi'^2(\tau_i)} \Delta t_i$, являющуюся интегральной суммой для функции $\sqrt{\varphi'^2(t) + \psi'^2(t)}$, непрерывной на отрезке $[\alpha, \beta]$ оси Ot . Пусть наибольшая из длин звеньев ломаной μ стремится к нулю. Тогда шаг разбиения отрезка $[\alpha, \beta]$ также стремится к нулю¹⁾. При этом интегральная сумма $\tilde{p}$ имеет предел, равный

$$\int_{\alpha}^{\beta} \sqrt{\varphi'^2(t) + \psi'^2(t)} dt.$$

Покажем, что длина s ломаной $M_0 M_1 \dots M_n$ при стремлении к нулю наибольшей из длин ее звеньев имеет тот же предел. Для этого достаточно доказать, что

$$\lim_{\mu \rightarrow 0} (\tilde{p} - p) = 0.$$

Оценим разность $\tilde{p} - p$ по абсолютной величине:

$$\begin{aligned} |\tilde{p} - p| &= \left| \sum_{i=1}^n [\sqrt{\varphi'^2(\tau_i) + \psi'^2(\tau_i)} - \sqrt{\varphi'^2(\tau_i) + \psi'^2(\tau_i^*)}] \Delta t_i \right| \leq \\ &\leq \sum_{i=1}^n |\sqrt{\varphi'^2(\tau_i) + \psi'^2(\tau_i)} - \sqrt{\varphi'^2(\tau_i) + \psi'^2(\tau_i^*)}| \Delta t_i \leq \\ &\leq \sum_{i=1}^n |\psi'(\tau_i) - \psi'(\tau_i^*)| \Delta t_i. \end{aligned}$$

Пусть $\varepsilon > 0$ — любое число. Так как функция $\psi'(t)$ на отрезке $[\alpha, \beta]$ непрерывна, то она равномерно непрерывна на этом отрезке, а значит, для числа $\frac{\varepsilon}{\beta - \alpha}$ существует число $\delta > 0$, такое, что для любых двух точек θ_1 и θ_2 отрезка $[\alpha, \beta]$, находящихся на расстоянии, меньшем δ , выполнено неравенство

$$|\psi'(\theta_1) - \psi'(\theta_2)| < \frac{\varepsilon}{\beta - \alpha}.$$

¹⁾ Доказательство этого утверждения читатель найдет в более полных курсах математического анализа (см., например, Фихтенгольц Г. М., Основы математического анализа, т. I).

²⁾ Мы воспользовались известным элементарным неравенством:

$$|\sqrt{a^2 + b^2} - \sqrt{a^2 + c^2}| \leq |b - c|,$$

которое можно доказать следующим образом:

$$\begin{aligned} |\sqrt{a^2 + b^2} - \sqrt{a^2 + c^2}| &= \frac{|b^2 - c^2|}{\sqrt{a^2 + b^2} + \sqrt{a^2 + c^2}} = \\ &= |b - c| \frac{|b + c|}{\sqrt{a^2 + b^2} + \sqrt{a^2 + c^2}} \leq |b - c| \frac{|b| + |c|}{\sqrt{a^2 + b^2} + \sqrt{a^2 + c^2}} \leq |b - c|, \end{aligned}$$

так как $|b| \leq \sqrt{a^2 + b^2}$, $|c| \leq \sqrt{a^2 + c^2}$.

Выбирая число μ столь малым, чтобы шаг соответствующего разбиения отрезка $[\alpha, \beta]$ на части был меньше δ , получаем:

$$|\tilde{p} - p| \leq \sum_{i=1}^n \frac{\varepsilon}{\beta - \alpha} \Delta t_i = \frac{\varepsilon}{\beta - \alpha} \sum_{i=1}^n \Delta t_i = \varepsilon.$$

Следовательно, $\lim_{\mu \rightarrow 0} p = \lim_{\mu \rightarrow 0} \tilde{p}$, т. е.

$$\lim_{\mu \rightarrow 0} p = \int_{\alpha}^{\beta} \sqrt{\varphi'^2(t) + \psi'^2(t)} dt.$$

Это означает, что кривая AB спрямляема и что ее длина s может быть найдена по формуле

$$s = \int_{\alpha}^{\beta} \sqrt{\varphi'^2(t) + \psi'^2(t)} dt \quad (2)$$

или короче по формуле

$$s = \int_{\alpha}^{\beta} \sqrt{x'^2 + y'^2} dt. \quad (2')$$

Примечание. Формула (2) верна и для замкнутой кривой, заданной уравнениями $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leq t \leq \beta$, при соответствующих условиях на функции $\varphi(t)$ и $\psi(t)$.

Действительно, фиксируя точку t_0 , $\alpha < t_0 < \beta$, разобьем соответствующей ей точкой M_0 кривую на две дуги AnM_0 и M_0mA (рис. 270). При этом

$$s_{AnM_0} = \int_{\alpha}^{t_0} \sqrt{x'^2 + y'^2} dt, \quad s_{M_0mA} = \int_{t_0}^{\beta} \sqrt{x'^2 + y'^2} dt,$$

$$s_{AnM_0mA} = \int_{\alpha}^{\beta} \sqrt{x'^2 + y'^2} dt.$$

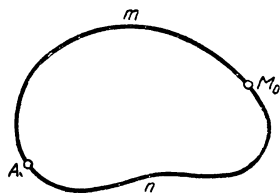


Рис. 270.

Пример 1. Найти длину дуги одной арки циклоиды $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $0 \leq t \leq 2\pi$. По формуле (2) имеем:

$$s = a \int_0^{2\pi} \sqrt{(1 - \cos t)^2 + \sin^2 t} dt = 2a \int_0^{2\pi} \sin \frac{t}{2} dt = 8a.$$

Пример 2. Найти длину эллипса $x = a \cos t$, $y = b \sin t$, $a > b$.

По формуле (2), применяемой на этот раз к вычислению длины замкнутой кривой, получаем:

$$s = \int_0^{2\pi} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt = \int_0^{2\pi} \sqrt{a^2 - (a^2 - b^2) \cos^2 t} dt = a \int_0^{2\pi} \sqrt{1 - \varepsilon^2 \cos^2 t} dt,$$

где $\varepsilon = \frac{\sqrt{a^2 - b^2}}{a}$ — эксцентриситет эллипса. Полученный интеграл в элементарных функциях не берется. Он называется *эллиптическим* интегралом второго рода.

Если кривая AB задана уравнением $y = f(x)$, $a \leq x \leq b$, где функция $f(x)$ непрерывна вместе со своей производной на отрезке $[a, b]$, то принимая x за параметр, из формулы (2) как ее частный случай получим формулу

$$s = \int_a^b \sqrt{1 + f'^2(x)} dx, \text{ или короче } s = \int_a^b \sqrt{1 + y'^2} dx. \quad (3)$$

Пример. Длина дуги полукубической параболы $y = x^{\frac{3}{2}}$ от ее вершины до точки $B(1, 1)$ вычисляется по формуле

$$s = \int_0^1 \sqrt{1 + \frac{9}{4}x} dx = \frac{8}{27} \left[\left(1 + \frac{9}{4}x\right)^{\frac{3}{2}} \right]_0^1 = \frac{13\sqrt{13}-8}{27}.$$

Пусть теперь кривая AB задана уравнением в полярных координатах:

$$\rho = f(\varphi), \quad \alpha \leq \varphi \leq \beta,$$

где $f(\varphi)$ — функция, имеющая для $\alpha \leq \varphi \leq \beta$ непрерывную производную.

Учитывая формулы $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, выражающие декартовы координаты точки через ее полярные координаты, получаем параметрические уравнения кривой AB с параметром φ в виде:

$$x = f(\varphi) \cos \varphi, \quad y = f(\varphi) \sin \varphi, \quad \alpha \leq \varphi \leq \beta.$$

При этом

$$\frac{dx}{d\varphi} = f'(\varphi) \cos \varphi - f(\varphi) \sin \varphi, \quad \frac{dy}{d\varphi} = f'(\varphi) \sin \varphi + f(\varphi) \cos \varphi,$$

$$\left(\frac{dx}{d\varphi}\right)^2 + \left(\frac{dy}{d\varphi}\right)^2 = f'^2(\varphi) + f^2(\varphi),$$

т. е. $s = \int_{\alpha}^{\beta} \sqrt{f'^2(\varphi) + f^2(\varphi)} d\varphi$ или короче

$$s = \int_{\alpha}^{\beta} \sqrt{\rho'^2 + \rho^2} d\varphi. \quad (4)$$

Пример. Найти длину дуги кардиоида $\rho = a(1 - \cos \varphi)$, $a > 0$. По формуле (4) имеем:

$$s = 2a \int_0^{\pi} \sqrt{(1 - \cos \varphi)^2 + \sin^2 \varphi} d\varphi = 4a \int_0^{\pi} \sin \frac{\varphi}{2} d\varphi = 8a.$$

3. Длина дуги пространственной кривой. Рассмотрим кривую AB , заданную параметрическими уравнениями:

$$x = \varphi(t), \quad y = \psi(t), \quad z = \chi(t), \quad \alpha \leq t \leq \beta,$$

таким образом, что при изменении параметра t от α к β соответствующая точка кривой перемещается в направлении от A к B . Тогда, если функции $\varphi(t)$, $\psi(t)$ и $\chi(t)$ непрерывны и имеют непрерывные производные на отрезке $[\alpha, \beta]$, то кривая AB спрямляема и ее длина s вычисляется по формуле

$$s = \int_{\alpha}^{\beta} \sqrt{x'^2 + y'^2 + z'^2} dt. \quad (5)$$

Доказательство этого утверждения аналогично рассмотренному выше доказательству формулы (2).

Пример. Найти длину дуги AB винтовой линии

$$x = \cos t, \quad y = \sin t, \quad z = 2t, \quad 0 \leq t \leq \pi.$$

По формуле (5):

$$s = \int_0^{\pi} \sqrt{\sin^2 t + \cos^2 t + 4} dt = \sqrt{5} \pi.$$

4. Дифференциал переменной длины дуги. Пусть AB — плоская кривая, заданная уравнениями $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leq t \leq \beta$, и удовлетворяющая высказанным при выводе формулы (2) условиям.

Возьмем на кривой AB точку M , соответствующую произвольному значению параметра t , $\alpha \leq t \leq \beta$. Длина дуги AM , вычисляемая, согласно (2), по формуле

$$s_{AM} = \int_{\alpha}^t \sqrt{x'^2 + y'^2} dt,$$

зависит от положения точки M , т. е. является функцией t , $\alpha \leq t \leq \beta$: $s_{AM} = s(t)$. Так как подинтегральная функция непрерывна на отрезке $[\alpha, \beta]$, то переменная дуга $s(t)$ имеет производную по t , равную, по теореме о производной определенного интеграла по его верхнему пределу, подинтегральной функции $\sqrt{x'^2 + y'^2}$ в точке t :

$$s'(t) = \sqrt{x'^2 + y'^2}. \quad (6)$$

Возводя это равенство в квадрат и умножая результат почленно на $(dt)^2$, получим формулу

$$(ds)^2 = (dx)^2 + (dy)^2 \quad (7)$$

для вычисления дифференциала длины дуги.

В частности, если кривая AB задана уравнением $y = f(x)$, $a \leq x \leq b^1$, то параметром является переменная x ,

$$s(x) = \int_a^x \sqrt{1 + y'^2} dx. \quad (8)$$

Формулы (6) и (7) записываются в виде:

$$s'(x) = \sqrt{1 + y'^2}, \quad (6')$$

$$(ds)^2 = (dx)^2 + (dy)^2. \quad (7')$$

Формула (7') допускает простое геометрическое истолкование: дифференциал длины дуги равен гипотенузе прямоугольного треугольника с катетами $|dx|$ и $|dy|$, т. е. равен длине отрезка касательной, проведенной к кривой в точке $M(x, y)$, от этой точки до точки с абсциссой $x + dx$ (рис. 271).

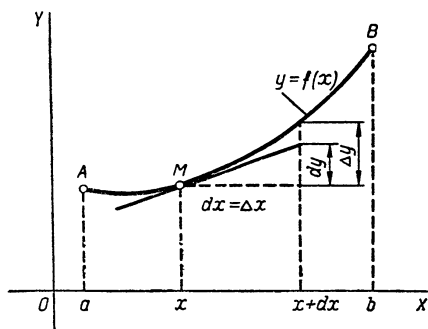


Рис. 271.

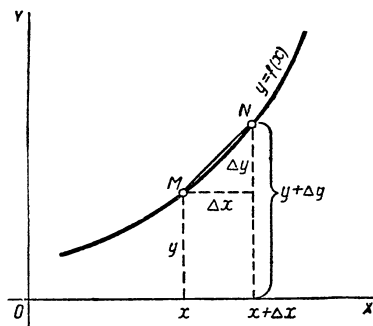


Рис. 272.

Используя выражение (8) длины дуги через интеграл, можно установить эквивалентность бесконечно малой дуги и стягивающей ее хорды.

Пусть $M(x, y)$ — любая фиксированная точка кривой AB , а $N(x + \Delta x, y + \Delta y)$ — переменная точка той же кривой (рис. 272).

¹⁾ Функция $f(x)$ предполагается непрерывной на отрезке $[a, b]$ вместе со своей производной.

Покажем, что при стремлении точки N к точке M отношение длины дуги MN к длине хорды MN (обозначим ее l_{MN}) стремится к единице:

$$\lim_{N \rightarrow M} \frac{s_{MN}}{l_{MN}} = 1.$$

Действительно,

$$\begin{aligned} s_{MN} &= |s_{AN} - s_{AM}| = \left| \int_a^{x+\Delta x} \sqrt{1+y'^2} dx - \int_a^x \sqrt{1+y'^2} dx \right| = \\ &= \left| \int_x^{x+\Delta x} \sqrt{1+f'^2(x)} dx \right|. \end{aligned}$$

Применяя теорему о среднем, получим:

$$s_{MN} = \sqrt{1+f'^2(\xi)} |\Delta x|,$$

где ξ — точка, лежащая на отрезке с концами x и $x + \Delta x$. С другой стороны,

$$l_{MN} = \sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2} |\Delta x|$$

или, по теореме Лагранжа,

$$l_{MN} = \sqrt{1+f'^2(\eta)} |\Delta x|,$$

где η — точка, лежащая между x и $x + \Delta x$.

Если точка N стремится к точке M , то Δx стремится к нулю.

В силу непрерывности функции $f'(x)$

$$\lim_{\Delta x \rightarrow 0} f'(\xi) = \lim_{\Delta x \rightarrow 0} f'(\eta) = f'(x).$$

Таким образом,

$$\lim_{N \rightarrow M} \frac{s_{MN}}{l_{MN}} = \lim_{\Delta x \rightarrow 0} \frac{\sqrt{1+f'^2(\xi)}}{\sqrt{1+f'^2(\eta)}} = \frac{\sqrt{1+f'^2(x)}}{\sqrt{1+f'^2(x)}} = 1.$$

В приближенных вычислениях для малых значений $dx = \Delta x$ вместо точного равенства (7') часто используются приближенные равенства (см. рис. 271, 272):

$$(\Delta s)^2 \approx (\Delta x)^2 + (\Delta y)^2 \quad \text{и} \quad (\Delta s)^2 \approx (dx)^2 + (dy)^2.$$

Примечание. Для пространственной кривой AB , рассмотренной в п. 3, длина переменной дуги AM вычисляется по формуле:

$$s(t) = \int_a^t \sqrt{x'^2 + y'^2 + z'^2} dt,$$

а для дифференциала длины дуги имеет место соотношение

$$ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}.$$

§ 13. Площадь поверхности вращения

Рассмотрим поверхность, образованную вращением кривой вокруг некоторой оси. Такая поверхность называется *поверхностью вращения*.

Площадь поверхности вращения кривой AB вокруг данной оси называют предел, к которому стремятся площади поверхностей вращения ломаных, вписанных в кривую AB , при стремлении к нулю наибольших из длин звеньев этих ломаных¹⁾.

¹⁾ Определение предела предоставляется дать читателю по аналогии с соответствующим определением предела в § 2.

При определенных условиях, наложенных на кривую AB , такой предел существует.

Покажем, например, что если лежащая в плоскости XOY кривая AB , заданная уравнением $y=f(x)$, $a \leq x \leq b$, где $f(x)$ — неотрицательная функция, непрерывная вместе со своей производной, вращается вокруг оси OX , то площадь поверхности вращения существует и вычисляется по формуле

$$S = 2\pi \int_a^b f(x) \sqrt{1 + f'^2(x)} dx \quad (1)$$

или короче:

$$S = 2\pi \int_a^b y \sqrt{1 + y'^2} dx. \quad (1')$$

Для доказательства этого утверждения разобьем дугу AB в направлении от A к B точками $M_0, M_1, \dots, M_n$ ¹⁾ на n частей и впишем в эту дугу ломаную $M_0M_1 \dots M_n$. Пусть $x_0 = a, x_1, \dots, x_n = b$ — абсциссы вершин этой ломаной ($x_0 < x_1 < \dots < x_n$). При вращении ломаной

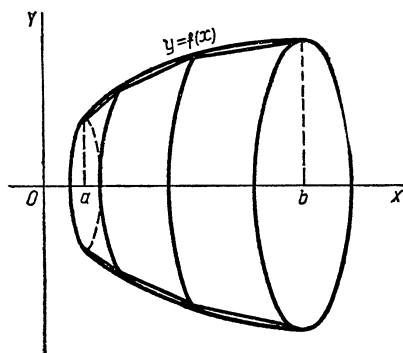


Рис. 273.

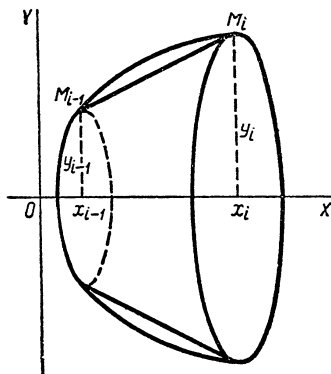


Рис. 274.

$M_0M_1 \dots M_n$ вокруг оси OX получится поверхность, состоящая из боковых поверхностей n усеченных конусов (рис. 273).

Рассмотрим i -й конус (рис. 274). Его высота равна Δx_i , радиусы оснований соответственно $y_{i-1} = f(x_{i-1})$ и $y_i = f(x_i)$, образующая $\sqrt{(x_i - x_{i-1})^2 + [f(x_i) - f(x_{i-1})]^2}$.

Площадь $\tilde{S}_i$ боковой поверхности этого конуса вычисляется по формуле

$$\tilde{S}_i = 2\pi \frac{f(x_{i-1}) + f(x_i)}{2} \sqrt{(x_i - x_{i-1})^2 + [f(x_i) - f(x_{i-1})]^2}.$$

Так как число $\frac{f(x_{i-1}) + f(x_i)}{2}$, как среднее арифметическое чисел $f(x_{i-1})$ и $f(x_i)$, заключено между ними, а функция $f(x)$ на отрезке $[a, b]$ непрерывна, то существует точка η_i , $x_{i-1} \leq \eta_i \leq x_i$, такая, что

$$f(\eta_i) = \frac{f(x_{i-1}) + f(x_i)}{2}.$$

Применяя затем к разности $f(x_i) - f(x_{i-1})$ теорему Лагранжа, получим:

$$f(x_i) - f(x_{i-1}) = f'(\xi_i)(x_i - x_{i-1}),$$

где ξ_i — точка, лежащая на отрезке $[x_{i-1}, x_i]$.

¹⁾ Здесь точка M_0 совпадает с A , а M_n — с B .

При этом

$$\tilde{S}_i = 2\pi f(\eta_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i, \quad \Delta x_i = x_i - x_{i-1},$$

а площадь $\tilde{S}$ поверхности, полученной в результате вращения ломаной $M_0 M_1 \dots M_n$, равна:

$$\tilde{S} = \sum_{i=1}^n 2\pi f(\eta_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i.$$

Сумма $\tilde{S}$ не является интегральной. Рассмотрим «близкую» к ней интегральную сумму:

$$\sum_{i=1}^n 2\pi f(\xi_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i$$

для функции

$$2\pi f(x) \sqrt{1 + f'^2(x)}.$$

Так как функции $f(x)$ и $f'(x)$ непрерывны на отрезке $[a, b]$, то функция $2\pi f(x) \sqrt{1 + f'^2(x)}$ также непрерывна на этом отрезке. При стремлении к нулю наибольшей из длин звеньев ломаной (обозначим ее μ) шаг разбиения отрезка $[a, b]$ стремится к нулю.

Переходя к пределу, получим:

$$\lim_{\mu \rightarrow 0} \sum_{i=1}^n 2\pi f(\xi_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i = 2\pi \int_a^b f(x) \sqrt{1 + f'^2(x)} dx.$$

Можно показать, что сумма $\tilde{S}$ стремится к тому же пределу.

Для этого достаточно доказать, что разность

$$\sum_{i=1}^n 2\pi f(\eta_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i - \sum_{i=1}^n 2\pi f(\xi_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i \quad (2)$$

стремится к нулю.

Так как функция $\sqrt{1 + f'^2(x)}$ непрерывна на отрезке $[a, b]$, то она ограничена на нем, т. е.

$$\sqrt{1 + f'^2(x)} < K, \quad a \leq x \leq b,$$

где K — некоторое число.

Пусть $\varepsilon > 0$ — любое число. Из равномерной непрерывности функции $f(x)$ на отрезке $[a, b]$ следует существование числа $\delta > 0$, такого, что для любых двух точек x' и x'' отрезка $[a, b]$, находящихся на расстоянии, меньшем δ , выполнено неравенство:

$$|f(x') - f(x'')| < \frac{\varepsilon}{2\pi K(b-a)}.$$

Выбирая длины звеньев ломаной столь мелкими, чтобы шаг разбиения отрезка $[a, b]$ был меньше δ , получим:

$$\begin{aligned} \left| \sum_{i=1}^n 2\pi f(\eta_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i - \sum_{i=1}^n 2\pi f(\xi_i) \sqrt{1 + f'^2(\xi_i)} \Delta x_i \right| &\leq \\ &\leq 2\pi \sum_{i=1}^n |f(\eta_i) - f(\xi_i)| \sqrt{1 + f'^2(\xi_i)} \Delta x_i < \\ &< 2\pi K \sum_{i=1}^n |f(\eta_i) - f(\xi_i)| \Delta x_i < 2\pi K \frac{\varepsilon}{2\pi K(b-a)} \sum_{i=1}^n \Delta x_i = \varepsilon. \end{aligned}$$

Так как $\varepsilon > 0$ — любое число, то отсюда следует, что разность (2) при $\mu \rightarrow 0$ стремится к нулю.

Таким образом, предел S площади поверхности вращения ломаной существует, причем

$$S = \lim_{\mu \rightarrow 0} \tilde{S} = 2\pi \int_a^b f(x) \sqrt{1 + f'^2(x)} dx.$$

Это означает, что существует площадь поверхности вращения кривой AB и для ее вычисления применима формула (1).

Пример. Найти площадь поверхности шарового пояса.

Рассмотрим окружность радиуса R с центром в начале координат. Вращая вокруг оси OX дугу $y = \sqrt{R^2 - x^2}$, $a \leq x \leq b$, этой окружности, получим поверхность шарового пояса. Ее площадь равна:

$$S = 2\pi \int_a^b \sqrt{R^2 - x^2} \sqrt{1 + \frac{x^2}{R^2 - x^2}} dx = 2\pi R (b - a).$$

Если кривая AB задана параметрическими уравнениями $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leq t \leq \beta$, где $\varphi(t)$ и $\psi(t)$ — функции, имеющие непрерывные производные, то, производя в определенном интеграле (1) замену переменной по формуле $x = \varphi(t)$, получим формулу для вычисления площади соответствующей поверхности вращения:

$$S = 2\pi \int_{\alpha}^{\beta} \psi(t) \sqrt{\varphi'^2(t) + \psi'^2(t)} dt. \quad (3)$$

Пример. Площадь поверхности, образованной при вращении вокруг оси OX одной арки циклоиды $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $0 \leq t \leq 2\pi$, равна:

$$\begin{aligned} S &= 2\pi \int_0^{2\pi} a(1 - \cos t) \sqrt{a^2(1 - \cos t)^2 + a^2 \sin^2 t} dt = \\ &= 2\pi a^2 \int_0^{2\pi} (1 - \cos t) \sqrt{2(1 - \cos t)} dt = \\ &= 8\pi a^2 \int_0^{2\pi} \sin^2 \frac{t}{2} dt = \\ &= 8\pi a^2 \int_0^{2\pi} \left(1 - \cos^2 \frac{t}{2}\right) \sin \frac{t}{2} dt = \\ &= 8\pi a^2 \left[-2 \cos \frac{t}{2} + \frac{2 \cos^3 \frac{t}{2}}{3} \right]_0^{2\pi} = \frac{64\pi a^2}{3}. \end{aligned}$$

Раздел 3

ПРИБЛИЖЕННЫЕ МЕТОДЫ ВЫЧИСЛЕНИЯ ОПРЕДЕЛЕННЫХ ИНТЕГРАЛОВ

§ 14. Приближенные методы вычисления определенных интегралов

Рассмотрим определенный интеграл $\int_a^b f(x) dx$ от функции $f(x)$, непрерывной на отрезке $[a, b]$.

Если функция $f(x)$ задана аналитически и ее первообразная на этом отрезке может быть найдена в элементарных функциях, то вычисление интеграла $\int_a^b f(x) dx$ сводится к применению формулы Ньютона — Лейбница. В тех случаях, когда функция $f(x)$ задана графически или таблично, а также в тех случаях, когда первообразная аналитически заданной функции $f(x)$ не выражается в элементарных функциях, для вычисления

ления интеграла $\int_a^b f(x) dx$ применяются приближенные методы, выражающие этот интеграл через значения функции $f(x)$, вычисленные для ряда значений независимой переменной x .

Основная идея приближенного вычисления интеграла $\int_a^b f(x) dx$ заключается в том, что подинтегральную функцию $f(x)$ заменяют другой, «близкой» к ней функцией, первообразная которой находится элементарным образом.

Пусть, например, известны значения функции $f(x)$ в точках $x_0, x_1, \dots, x_m$ отрезка $[a, b]$.

Возьмем в качестве функции, «близкой» к функции $f(x)$, многочлен $P(x)$ степени не выше m , совпадающий с $f(x)$ в точках $x_0, x_1, \dots, x_m$. Этот интерполяционный многочлен можно записать в виде (см. гл. IV, § 29):

$$P(x) = \sum_{i=0}^m f(x_i) \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_m)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_m)}.$$

Полагая, что $f(x) \approx P(x)$, приходим к следующей формуле приближенного интегрирования:

$$\int_a^b f(x) dx \approx \int_a^b P(x) dx. \quad (1)$$

В случае, если функция $f(x)$ на отрезке $[a, b]$ имеет непрерывные производные до порядка $m+1$ включительно, погрешность формулы (1) легко оценивается с помощью остаточного члена интерполяционной формулы Лагранжа (гл. IV, § 29):

$$\begin{aligned} r &= \left| \int_a^b f(x) dx - \int_a^b P(x) dx \right| = \left| \int_a^b [f(x) - P(x)] dx \right| = \\ &= \left| \int_a^b \frac{f^{(m+1)}(\xi)}{(m+1)!} (x-x_0)(x-x_1)\dots(x-x_m) dx \right| \leq \\ &\leq \frac{M_{m+1}}{(m+1)!} \int_a^b |(x-x_0)(x-x_1)\dots(x-x_m)| dx. \end{aligned}$$

Здесь ξ — промежуточная точка отрезка $[a, b]$, а

$$M_{m+1} = \max_{[a, b]} |f^{(m+1)}(x)|.$$

В частности, если $m=0$ и $x_0=a$, то $P(x) = f(a)$,

$$f(x) \approx f(a), \quad \int_a^b f(x) dx \approx f(a)(b-a), \quad (2)$$

причем

$$r \leq \frac{M_1}{2} (b-a)^2, \quad (3)$$

где

$$M_1 = \max_{[a, b]} |f'(x)|.$$

В случае, если функция $f(x)$ неотрицательна на отрезке $[a, b]$, равенство (2) геометрически означает, что площадь криволинейной трапеции с основанием $[a, b]$, ограниченной сверху кривой $y=f(x)$, прибли-

женно принята равной площади прямоугольника с тем же основанием и высотой $f(a)$ (рис. 275).

В связи с этим формула (2) называется *элементарной формулой прямоугольников*.

Полагая $m=1$, $x_0=a$, $x_1=b$, получаем интерполяционный многочлен $P(x)$ первой степени:

$$P(x) = f(a) \frac{x-b}{a-b} + f(b) \frac{x-a}{b-a},$$

его график — прямая, проходящая через точки $(a, f(a))$ и $(b, f(b))$.

При этом

$$\begin{aligned} \int_a^b f(x) dx &\approx \int_a^b P(x) dx = \frac{f(a)}{a-b} \int_a^b (x-b) dx + \\ &+ \frac{f(b)}{b-a} \int_a^b (x-a) dx = \frac{b-a}{2} [f(a) + f(b)], \end{aligned}$$

т. е.

$$\int_a^b f(x) dx \approx \frac{b-a}{2} [f(a) + f(b)]. \quad (4)$$

Погрешность формулы (4) оценивается неравенством

$$r \leq \frac{M_2}{2!} \int_a^b |(x-a)(x-b)| dx = -\frac{M_2}{2!} \int_a^b (x-a)(b-x) dx = \frac{M_2}{12} (b-a)^3,$$

где $M_2 = \max_{[a,b]} |f''(x)|$.

Если функция $f(x)$ неотрицательна на отрезке $[a, b]$, то равенство (4) геометрически означает, что площадь криволинейной трапеции с основа-

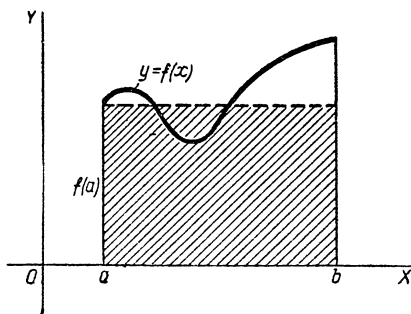


Рис. 275.

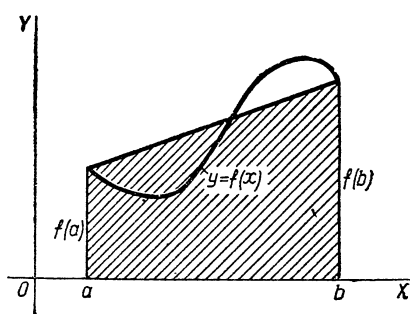


Рис. 276.

нием $[a, b]$, ограниченной сверху кривой $y=f(x)$, приближенно принята равной площади прямолинейной трапеции с основаниями $f(a)$ и $f(b)$ и высотой $b-a$ (рис. 276).

Формула (4) носит название *элементарной формулы трапеций*.

Если $m=2$, $x_0=a$, $x_1=\frac{a+b}{2}$, $x_2=b$, то интерполяционный многочлен имеет вторую степень:

$$\begin{aligned} P(x) = f(a) \frac{\left(x - \frac{a+b}{2}\right)(x-b)}{\left(a - \frac{a+b}{2}\right)(a-b)} + f\left(\frac{a+b}{2}\right) \frac{(x-a)(x-b)}{\left(\frac{a+b}{2} - a\right)\left(\frac{a+b}{2} - b\right)} + \\ + f(b) \frac{(x-a)\left(x - \frac{a+b}{2}\right)}{(b-a)\left(b - \frac{a+b}{2}\right)}. \end{aligned}$$

Графиком такого многочлена является парабола второй степени с осью, параллельной оси OY , проходящая через точки

$$(a, f(a)), \left(\frac{a+b}{2}, f\left(\frac{a+b}{2}\right) \right), (b, f(b)).$$

При этом

$$\begin{aligned} \int_a^b f(x) dx &\approx \int_a^b P(x) dx = \\ &= \frac{2f(a)}{(a-b)^2} \int_a^b \left(x - \frac{a+b}{2}\right)(x-b) dx + \frac{4f\left(\frac{a+b}{2}\right)}{(a-b)(b-a)} \int_a^b (x-a)(x-b) dx + \\ &+ \frac{2f(b)}{(b-a)^2} \int_a^b (x-a)\left(x - \frac{a+b}{2}\right) dx. \end{aligned}$$

Произведя замену переменной $x = a + t(b-a)$, получим:

$$\begin{aligned} \int_a^b f(x) dx &\approx 2f(a)(b-a) \int_0^1 \left(t - \frac{1}{2}\right)(t-1) dt - \\ &- 4f\left(\frac{a+b}{2}\right)(b-a) \int_0^1 t(t-1) dt + 2f(b)(b-a) \int_0^1 t\left(t - \frac{1}{2}\right) dt = \\ &= \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]. \end{aligned}$$

Формула

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \quad (5)$$

называется *элементарной формулой парабол*.

Оценка погрешности формулы (5) дается неравенством

$$\begin{aligned} r &\leq \frac{M_3}{3!} \int_a^b \left| (x-a) \left(x - \frac{a+b}{2}\right) (x-b) \right| dx = \\ &= \frac{M_3}{6} (b-a)^4 \int_0^1 \left| t \left(t - \frac{1}{2}\right) (t-1) \right| dt = \frac{M_3}{192} (b-a)^4, \end{aligned}$$

где

$$M_3 = \max_{[a, b]} |f'''(x)|.$$

Для функции $f(x)$, неотрицательной на отрезке $[a, b]$, формула (5) геометрически означает, что площадь криволинейной трапеции с основанием $[a, b]$, ограниченной сверху кривой $y=f(x)$, заменена площадью криволинейной трапеции с тем же основанием, ограниченной сверху параболой $y=P(x)$ (рис. 277).

Элементарные формулы прямоугольников, трапеций и парабол дают лишь грубое приближение интеграла $\int_a^b f(x) dx$.

Для получения более точных формул разбивают отрезок $[a, b]$ на части, интеграл $\int_a^b f(x) dx$ заменяют суммой интегралов по отрезкам деления, а каждый из полученных интегралов вычисляют приближенно по одной из указанных выше формул.

Общая формула прямоугольников. Разобьем отрезок $[a, b]$ точками $x_0=a, x_1, \dots, x_n=b$ ($x_0 < x_1 < \dots < x_n=b$) на n равных частей.

При этом

$$\int_a^b f(x) dx = \int_a^{x_1} f(x) dx + \int_{x_1}^{x_2} f(x) dx + \dots + \int_{x_{n-1}}^{x_n} f(x) dx.$$

Применяя для вычисления каждого из интегралов в правой части равенства формулу (2), получим приближенное равенство:

$$\int_a^b f(x) dx \approx \frac{b-a}{n} [y_0 + y_1 + \dots + y_{n-1}], \quad (6)$$

где

$$y_0 = f(x_0), y_1 = f(x_1), \dots, y_{n-1} = f(x_{n-1}), \text{ а } \frac{b-a}{n}$$

есть длина каждого из отрезков деления.

Эта формула называется *общей формулой прямоугольников*.

В случае, если $f(x) \geq 0$ на отрезке $[a, b]$, формула (6) геометрически означает замену площади криволинейной трапеции с основанием $[a, b]$, ограниченной сверху кривой $y = f(x)$, площадью ступенчатой фигуры, составленной из прямоугольников (рис. 278).

Оценка погрешности в формуле (6) дается неравенством

$$\begin{aligned} R &= \left| \int_a^b f(x) dx - \frac{b-a}{n} [y_0 + y_1 + \dots + y_{n-1}] \right| \leq \\ &\leq n \cdot \frac{M_1}{2} \left(\frac{b-a}{n} \right)^2 = \frac{M_1}{2n} (b-a)^2, \end{aligned}$$

где $M_1 = \max_{[a, b]} |f'(x)|$.

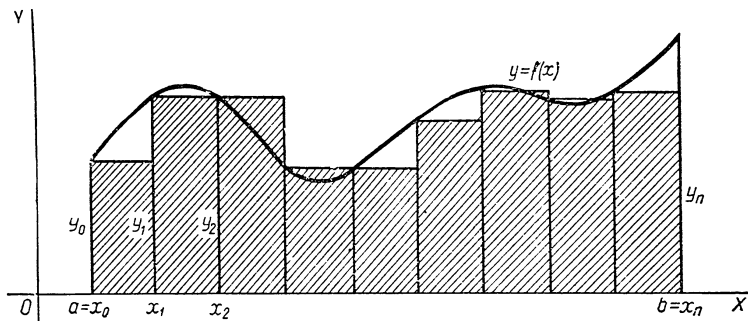


Рис. 278.

Общая формула трапеций. Разобьем отрезок $[a, b]$ точками $x_0 = a, x_1, \dots, x_n = b$ ($x_0 < x_1 < \dots < x_n$) на n равных частей и обозначим соответственно через $y_0, y_1, \dots, y_n$ значения функции $f(x)$ в точках деления. Длины отрезков деления равны $\frac{b-a}{n}$.

По свойствам определенного интеграла

$$\int_a^b f(x) dx = \int_a^{x_1} f(x) dx + \int_{x_1}^{x_2} f(x) dx + \dots + \int_{x_{n-1}}^{x_n} f(x) dx.$$

Применяя к каждому из интегралов в правой части равенства формулу (4), получим приближенную формулу:

$$\int_a^b f(x) dx \approx \frac{b-a}{2n} [y_0 + y_n + 2(y_1 + y_2 + \dots + y_{n-1})], \quad (7)$$

называемую *общей формулой трапеций*.

Погрешность формулы трапеций оценивается следующим образом:

$$\begin{aligned} R &= \left| \int_a^b f(x) dx - \frac{b-a}{2n} [y_0 + y_n + 2(y_1 + y_2 + \dots + y_{n-1})] \right| \leq \\ &\leq n \frac{M_2}{12} \left(\frac{b-a}{n} \right)^3 = \frac{M_2}{12n^2} (b-a)^3, \end{aligned}$$

где $M_2 = \max_{[a,b]} |f''(x)|$.

Геометрически вычисление интеграла $\int_a^b f(x) dx$ от неотрицательной функции по формуле трапеций сводится к замене площади соответствующей

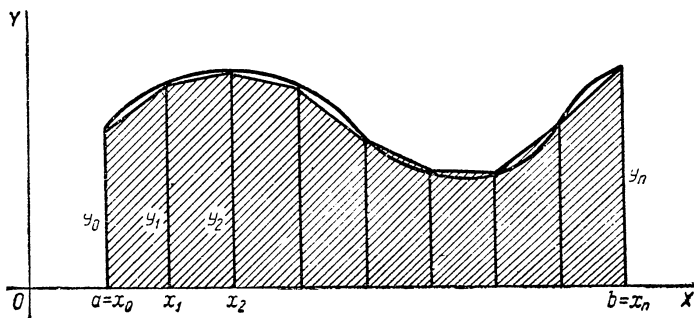


Рис. 279.

ющей криволинейной трапеции площадью, составленной из прямолинейных трапеций (рис. 279).

Общая формула парабол (формула Симпсона). Разобьем отрезок $[a, b]$ точками $x_0 = a, x_1, \dots, x_{2n} = b, x_0 < x_1 < \dots < x_{2n}$, на $2n$ равных частей и обозначим через $y_0, y_1, \dots, y_{2n}$ значения функции $f(x)$ в точках деления. При этом

$$\int_a^b f(x) dx = \int_{x_0}^{x_2} f(x) dx + \int_{x_2}^{x_4} f(x) dx + \dots + \int_{x_{2n-2}}^{x_{2n}} f(x) dx.$$

Применяя к каждому из полученных интегралов элементарную формулу парабол (5), получим приближенную формулу:

$$\begin{aligned} \int_a^b f(x) dx &\approx \frac{b-a}{6n} \left[y_0 + y_{2n} + 4(y_1 + y_3 + \dots + y_{2n-1}) + \right. \\ &\quad \left. + 2(y_2 + y_4 + \dots + y_{2n-2}) \right], \end{aligned}$$

называемую *общей формулой парабол* (см. рис. 280).

Погрешность общей формулы парабол дается неравенством

$$\begin{aligned} R &= \left| \int_a^b f(x) dx - \frac{b-a}{6n} \left[y_0 + y_{2n} + 4(y_1 + y_3 + \dots + y_{2n-1}) + \right. \right. \\ &\quad \left. \left. + 2(y_2 + y_4 + \dots + y_{2n-2}) \right] \right| \leq n \frac{M_3}{192} \left(\frac{b-a}{n} \right)^4 = \frac{M_3}{192n^3} (b-a)^4, \end{aligned}$$

где

$$M_3 = \max_{[a, b]} |f'''(x)|.$$

Примечание. Если функция $f(x)$ имеет на отрезке $[a, b]$ непрерывную производную четвертого порядка, то, как показывают более де-

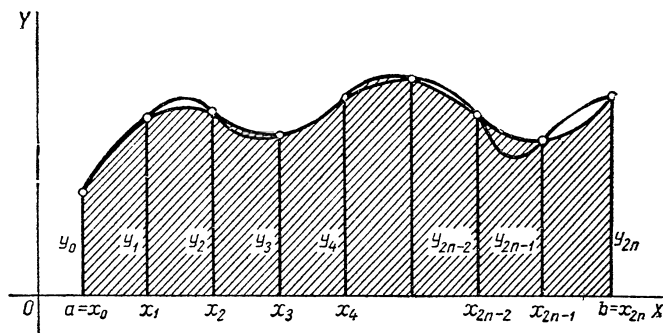


Рис. 280.

тальные исследования, погрешность общей формулы парабол можно оценить неравенством:

$$R \leq \frac{M_4}{2880n^4} (b-a)^5, \text{ где } M_4 = \max_{[a, b]} |f^{IV}(x)|.$$

Пример. Вычислим по формуле трапеций интеграл

$$\int_0^1 e^{-x^2} dx,$$

взяв $n = 5$. Так как

$$f(x) = e^{-x^2}, \text{ то } f'(x) = -2xe^{-x^2}, f''(x) = 2(2x^2 - 1)e^{-x^2}.$$

На отрезке $[0, 1]$ имеем $|f''(x)| \leq 2$. Так как $R \leq \frac{M_2}{12n^2} (b-a)^3$, то имеем $R \leq \frac{2}{12 \cdot 25} < \frac{1}{100}$.

Следовательно, при вычислении интеграла $\int_0^1 e^{-x^2} dx$ по формуле трапеций значения функции e^{-x^2} в точках деления нужно вычислять не более, чем с тремя десятичными знаками.

Вычисления располагаем в таблицу:

N	x	$-x^2$	e^{-x^2}
0	0,0	—0,00	1,000
1	0,2	—0,04	0,961
2	0,4	—0,16	0,852
3	0,6	—0,36	0,698
4	0,8	—0,64	0,527
5	1,0	—1,00	0,368

$$\begin{aligned} \int_0^1 e^{-x^2} dx &\approx \frac{1}{10} [1,000 + 0,368 + \\ &+ 2(0,961 + 0,852 + 0,698 + 0,527)] = \\ &= \frac{1}{10} [1,368 + 2 \cdot 3,038] = 0,744^1. \end{aligned}$$

В результате получаем $\int_0^1 e^{-x^2} dx \approx 0,74$ с точностью до 0,01.

¹⁾ При таких вычислениях ошибка от округления значений в точках x_1, x_2, x_3, x_4, x_5 в сумме не превосходит 0,005.

Поставим теперь задачу: вычислить интеграл $\int_0^1 e^{-x^2} dx$ по формуле парабол с точностью до 0,001.

Так как $f^{IV}(x) = 4e^{-x^2}(4x^4 - 12x^2 + 3)$, то на отрезке $[0, 1]$ $|f^{IV}(x)| \leq 20$ (так как $e^{-x^2} \leq 1$, $|4x^4 - 12x^2 + 3| \leq 5$). Используя оценку $R \leq \frac{M_4}{2880n^4} (b-a)^5$, получаем:

$$\frac{20}{2880n^4} < \frac{1}{1000}, \quad n^4 > \frac{1000}{144}.$$

Для выполнения этого неравенства достаточно взять $n=2$, т. е. $2n=4$.

N	x	$-x^2$	e^{-x^2}
0	0,00	-0,0000	1,0000
1	0,25	-0,0625	0,9394
2	0,50	-0,2500	0,7788
3	0,75	-0,5625	0,5698
4	1,00	-1,0000	0,3679

$$\int_0^1 e^{-x^2} dx = \frac{1}{12} [1,0000 + 0,3679 + 4(0,9394 + 0,5698) + 2 \cdot 0,7788] \approx 0,7469.$$

Таким образом, с точностью до 0,001

$$\int_0^1 e^{-x^2} dx \approx 0,747.$$

Примечание. Если функция $f(x)$ имеет на отрезке $[a, b]$ непрерывные производные до четвертого порядка включительно, то, как следует из полученных оценок точности формул приближенного интегрирования, наиболее точные результаты при одном и том же значении n дает, вообще говоря, формула парабол, а наиболее грубое приближение — формула прямоугольников.

Раздел 4

ОБОБЩЕНИЕ ПОНЯТИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

§ 15. Несобственные интегралы

Вводя определенный интеграл как предел интегральных сумм, мы предполагали, что отрезок интегрирования конечный, а подинтегральная функция ограничена на этом отрезке. Если хоть одно из этих условий не выполнено, данное выше определение теряет смысл. Рассмотрим некоторые возможные обобщения понятия определенного интеграла.

1. Несобственные интегралы с бесконечными пределами интегрирования.

Пусть функция $f(x)$ непрерывна на промежутке $a \leq x < +\infty$, тогда она непрерывна на любом отрезке $[a, b]$, $b \geq a$, а следовательно, существует интеграл

$$\int_a^b f(x) dx.$$

Этот интеграл является функцией своего верхнего предела b , определенной на промежутке $a \leq b < +\infty$. Если при $b \rightarrow +\infty$ интеграл $\int_a^b f(x) dx$ стремится к конечному пределу, то этот предел обозначают

символом $\int_a^{+\infty} f(x) dx$ и называют *несобственным интегралом* по бесконечному промежутку (от a до бесконечности) от функции $f(x)$. Таким образом, по определению $\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} \int_a^b f(x) dx$, если этот предел существует и конечен. В этом случае принято говорить, что несобственный интеграл $\int_a^{+\infty} f(x) dx$ *существует* или *сходится*.

Если же при $b \rightarrow +\infty$ не существует предела интеграла $\int_a^b f(x) dx$ или этот предел бесконечен, то символу $\int_a^{+\infty} f(x) dx$ никакого числового смысла не приписывают и, называя его снова *несобственным интегралом*, говорят, что этот несобственный интеграл *не существует* или *расходится*.

Аналогичным образом для функции $f(x)$, непрерывной на бесконечном промежутке $-\infty < x \leq b$, определяется *несобственный интеграл* $\int_{-\infty}^b f(x) dx$.

При этом

$$\int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx,$$

если этот предел существует и конечен.

Для функции $f(x)$, непрерывной на всей числовой оси, *несобственный интеграл* $\int_{-\infty}^{+\infty} f(x) dx$ определяется равенством:

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^c f(x) dx + \int_c^{+\infty} f(x) dx, \quad (1)$$

где c — любое число, если каждый из интегралов в правой части равенства (1) сходится. При этом несобственный интеграл $\int_{-\infty}^{+\infty} f(x) dx$ называется *сходящимся*. Если хотя бы один из интегралов $\int_{-\infty}^c f(x) dx$ и $\int_c^{+\infty} f(x) dx$ расходится, то несобственный интеграл $\int_{-\infty}^{+\infty} f(x) dx$ называется *расходящимся*.

Из данных выше определений сходящихся несобственных интегралов с бесконечными пределами интегрирования следует, что эти интегралы являются не пределами интегральных сумм, а пределами определенных интегралов с переменными верхними или нижними пределами при стремлении этих пределов к бесконечности.

Если функция $f(x)$ непрерывна и неотрицательна на бесконечном промежутке $[a, +\infty)$, то $\int_a^{+\infty} f(x) dx$ численно равен площади криволинейной трапеции, ограниченной снизу отрезком $[a, b]$ оси OX , сверху — кривой $y = f(x)$, слева и справа — прямыми $x = a$ и $x = b$. При возрастании b прямая $x = b$, ограничивающая эту криволинейную трапецию, двигается направо.

В случае, если при этом несобственный интеграл $\int_a^{+\infty} f(x) dx$ сходится, его величину естественно принять за площадь бесконечной полосы,

ограниченной снизу осью OX , сверху — графиком функции $y=f(x)$, слева — прямой $x=a$ (рис. 281).

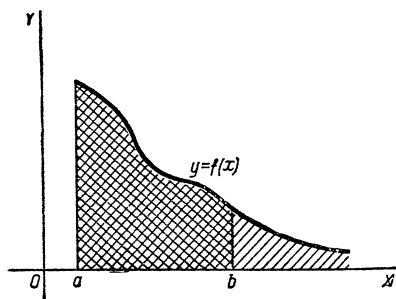


Рис. 281.

Аналогичные рассуждения имеют место для интегралов вида:

$$\int_{-\infty}^b f(x) dx \text{ и } \int_{-\infty}^{+\infty} f(x) dx.$$

Рассмотрим несколько примеров несобственных интегралов.

Пример 1. $\int_0^{+\infty} \frac{dx}{1+x^2} = \lim_{b \rightarrow +\infty} \int_0^b \frac{dx}{1+x^2} = \lim_{b \rightarrow +\infty} \arctg x \Big|_0^b = \lim_{b \rightarrow +\infty} \arctg b = \frac{\pi}{2}.$

Пример 2. $\int_0^{+\infty} \cos x dx$ расходится, так как

$$\int_0^b \cos x dx = \sin x \Big|_0^b = \sin b.$$

Предела у функции $\sin b$ при $b \rightarrow +\infty$ не существует.

Пример 3. $\int_{-\infty}^1 e^{2x} dx = \lim_{a \rightarrow -\infty} \int_a^1 e^{2x} dx = \lim_{a \rightarrow -\infty} \frac{1}{2} e^{2x} \Big|_a^1 =$
 $= \frac{1}{2} \lim_{a \rightarrow -\infty} (e^2 - e^{2a}) = \frac{1}{2} e^2.$

Пример 4. $\int_{-\infty}^{+\infty} e^x dx$ расходится, так как несобственный интеграл $\int_0^{+\infty} e^x dx$

расходится: $\lim_{b \rightarrow +\infty} \int_0^b e^x dx = \lim_{b \rightarrow +\infty} e^x \Big|_0^b = \lim_{b \rightarrow +\infty} (e^b - 1) = \infty.$

Пример 5. Исследуем на сходимость интеграл вида $\int_1^{+\infty} \frac{dx}{x^\alpha}$, где α — некоторое число.

1. Если $\alpha \neq 1$, то для любого $b > 0$

$$\int_1^b \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} (b^{1-\alpha} - 1),$$

$$\lim_{b \rightarrow +\infty} \int_1^b \frac{dx}{x^\alpha} = \begin{cases} \frac{-1}{1-\alpha}, & \alpha > 1, \\ \infty, & \alpha < 1. \end{cases}$$

2. Если $\alpha = 1$, то для любого $b > 0$

$$\int_1^b \frac{dx}{x} = \ln b, \quad \lim_{b \rightarrow +\infty} \int_1^b \frac{dx}{x} = \infty.$$

Таким образом, несобственный интеграл $\int_1^{+\infty} \frac{dx}{x^\alpha}$ сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$.

Определенные выше несобственные интегралы обладают рядом свойств, присущих обыкновенным определенным интегралам. В частности, для них можно получить обобщение формулы Ньютона — Лейбница.

Пусть, например, функция $f(x)$ непрерывна на промежутке $[a, +\infty)$ и имеет на этом промежутке первообразную $F(x)$. Тогда для любого b , $a \leq b < +\infty$, по формуле Ньютона — Лейбница

$$\int_a^b f(x) dx = F(b) - F(a) = F(x) \Big|_a^b.$$

Отсюда следует, что несобственный интеграл $\int_a^{+\infty} f(x) dx$ существует в том, и только в том, случае, когда существует конечный предел $F(b)$ при $b \rightarrow +\infty$.

Если этот предел существует, то

$$\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} F(b) - F(a).$$

Условимся сокращенно обозначать $\lim_{b \rightarrow +\infty} F(b)$ как $F(+\infty)$.

Тогда

$$\int_a^{+\infty} f(x) dx = F(+\infty) - F(a) = F(x) \Big|_a^{+\infty}.$$

Эту формулу можно рассматривать как обобщение формулы Ньютона — Лейбница на случай несобственного интеграла.

Аналогично при выполнении соответствующих условий получаются формулы

$$\int_{-\infty}^b f(x) dx = F(x) \Big|_{-\infty}^b$$

и

$$\int_{-\infty}^{+\infty} f(x) dx = F(x) \Big|_{-\infty}^{+\infty},$$

если под $F(-\infty)$ понимать $\lim_{a \rightarrow -\infty} F(a)$.

Признаки сходимости несобственных интегралов будут рассматриваться в гл. X, § 3.

II. Несобственные интегралы от неограниченных функций. Пусть теперь функция $f(x)$ непрерывна на полуинтервале $[a, b)$ и не ограничена вблизи b .

Эта функция непрерывна, а значит, интегрируема на любом отрезке $[a, b_1]$, $a \leq b_1 < b$.

Интеграл $\int_a^{b_1} f(x) dx$ будет функцией своего верхнего предела, определенной на промежутке

$$a \leq b_1 < b.$$

Если при $b_1 \rightarrow b - 0$ существует конечный предел интеграла $\int_a^{b_1} f(x) dx$,

то этот предел обозначают символом $\int_a^b f(x) dx$ и называют *несобственным интегралом* от неограниченной функции $f(x)$ по отрезку $[a, b]$.

Таким образом,

$$\int_a^b f(x) dx = \lim_{b_1 \rightarrow b-0} \int_a^{b_1} f(x) dx,$$

если этот предел существует и конечен. В этом случае принято говорить, что несобственный интеграл $\int_a^b f(x) dx$ *сходится*.

Если же при $b_1 \rightarrow b - 0$ предела интеграла $\int_a^{b_1} f(x) dx$ не существует или он бесконечен, то символу $\int_a^b f(x) dx$, который снова называют *несобственным интегралом*, никакого числового смысла не приписывают, говоря при этом, что несобственный интеграл $\int_a^b f(x) dx$ *не существует* или *расходится*.

Аналогично для функции $f(x)$, непрерывной на полуинтервале $(a, b]$ и неограниченной вблизи a , вводится понятие *несобственного интеграла* $\int_a^b f(x) dx$.

Полагают, что

$$\int_a^b f(x) dx = \lim_{a_1 \rightarrow a+0} \int_{a_1}^b f(x) dx,$$

если этот предел существует и конечен.

Несобственный интеграл от функции $f(x)$, непрерывной на интервале (a, b) и неограниченной вблизи его концов a и b , определяется равенством

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx,$$

где c — любая точка интервала (a, b) , если каждый из несобственных интегралов $\int_a^c f(x) dx$ и $\int_c^b f(x) dx$ сходится. При этом несобственный интеграл $\int_a^b f(x) dx$ называется *сходящимся*. Если хотя бы один из интегралов $\int_a^c f(x) dx$ и $\int_c^b f(x) dx$ расходится, то несобственный интеграл $\int_a^b f(x) dx$ называют *расходящимся*.

Пусть теперь функция $f(x)$ непрерывна на отрезке $[a, b]$ всюду, кроме некоторой точки c , $a < c < b$, и не ограничена вблизи c . Несобственный интеграл $\int_a^b f(x) dx$ определяется равенством

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx,$$

если каждый из интегралов в правой части равенства сходится. Если хотя бы один из этих интегралов расходится, то несобственный интеграл $\int_a^b f(x) dx$ называется *расходящимся*.

Аналогичным образом определяется *несобственный интеграл* по отрезку $[a, b]$ от функции, непрерывной на нем всюду, кроме конечного числа точек, и неограниченной вблизи этих точек.

Несобственные интегралы от неограниченных функций по конечному отрезку $[a, b]$, так же как и несобственные интегралы с бесконечными пределами, являются не пределами интегральных сумм, а пределами определенных интегралов с переменными верхними или нижними пределами.

Если функция $f(x)$ непрерывна и неотрицательна на полуинтервале $[a, b)$, не ограничена вблизи b и несобственный интеграл $\int_a^b f(x) dx$ сходится, то его вели-

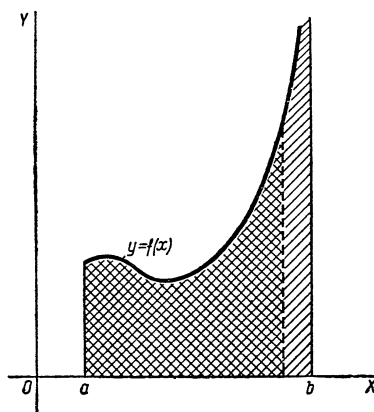


Рис. 282.

чину можно принять за площадь бесконечной полосы, ограниченной снизу отрезком $[a, b]$ оси OX , слева и справа — прямыми $x=a$, $x=b$, сверху — кривой $y=f(x)$ (см. рис. 282). Аналогично — для несобственных интегралов другого типа.

Пример 1. Интеграл $\int_0^1 \frac{dx}{x^\alpha}$, где α — положительное число, несобственный.

Для того чтобы исследовать его на сходимость, рассмотрим интеграл $\int_{a_1}^1 \frac{dx}{x^\alpha}$, $0 < a_1 \leq 1$. Если $\alpha \neq 1$, то

$$\int_{a_1}^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} x^{1-\alpha} \Big|_{a_1}^1 = \frac{1}{1-\alpha} (1 - a_1^{1-\alpha}).$$

Если $\alpha = 1$, то

$$\int_{a_1}^1 \frac{dx}{x} = \ln x \Big|_{a_1}^1 = -\ln a_1.$$

Переходя к пределу при $a_1 \rightarrow 0+0$, получаем:

$$\lim_{a_1 \rightarrow 0+0} \int_{a_1}^1 \frac{dx}{x^\alpha} = \begin{cases} \infty, & \alpha \geq 1, \\ \frac{1}{1-\alpha}, & \alpha < 1. \end{cases}$$

Таким образом, несобственный интеграл $\int_0^1 \frac{dx}{x^\alpha}$ сходится при $\alpha < 1$ и расходится при $\alpha \geq 1$.

Ряд свойств, присущих определенным интегралам, распространяется и на несобственные интегралы от неограниченных функций. В частности, для этих несобственных интегралов имеет место обобщение формулы Ньютона — Лейбница.

Пусть, например, функция $f(x)$ непрерывна на полуинтервале $[a, b)$ и не ограничена вблизи b . Обозначая через $F(x)$ первообразную функции

$f(x)$ на полуинтервале $[a, b)$, по формуле Ньютона — Лейбница получаем:

$$\int_a^{b_1} f(x) dx = F(b_1) - F(a) = F(x) \Big|_a^{b_1}$$

для любого b_1 , $a \leq b_1 < b$. Отсюда следует, что несобственный интеграл $\int_a^b f(x) dx$ существует в том, и только в том, случае, когда существует конечный предел $F(b_1)$ при $b_1 \rightarrow b - 0$. Если этот предел существует, то

$$\int_a^b f(x) dx = \lim_{b_1 \rightarrow b-0} F(b_1) - F(a).$$

Полагая $F(b)$ равным пределу $F(b_1)$ при $b_1 \rightarrow b - 0$, т. е. доопределяя функцию $F(x)$ в точке b по непрерывности слева, будем иметь:

$$\int_a^b f(x) dx = F(b) - F(a) = F(x) \Big|_a^b.$$

Эта формула является *обобщением формулы Ньютона — Лейбница* на случай несобственного интеграла указанного вида.

Аналогично, если функция $f(x)$ непрерывна на полуинтервале $(a, b]$ и не ограничена вблизи a [или непрерывна на интервале (a, b) и не ограничена вблизи a и $b]$, то

$$\int_a^b f(x) dx = F(x) \Big|_a^b,$$

где $F(x)$ — первообразная функции $f(x)$ на полуинтервале $(a, b]$ [на интервале (a, b)] и под $F(a)$ понимается предел $F(x)$ при $x \rightarrow a + 0$.

Пусть теперь функция $f(x)$ непрерывна на отрезке $[a, b]$ всюду, кроме точек $c_1, c_2, \dots, c_m$, и не ограничена вблизи этих точек.

Если существует функция $F(x)$, непрерывная на всем отрезке $[a, b]$ и являющаяся первообразной для функции $f(x)$ на промежутках $[a, c_1)$, (c_1, c_2) , $\dots$, $(c_m, b]$, то несобственный интеграл $\int_a^b f(x) dx$ сходится и его вычисление можно производить по формуле

$$\int_a^b f(x) dx = F(x) \Big|_a^b,$$

называемой *формулой Ньютона — Лейбница*.

Действительно, каждый из интегралов

$$\int_a^{c_1} f(x) dx, \quad \int_{c_1}^{c_2} f(x) dx, \quad \dots, \quad \int_{c_m}^b f(x) dx$$

в данных условиях сходится и может быть вычислен по формуле Ньютона — Лейбница. Следовательно интеграл $\int_a^b f(x) dx$ сходится, причем

$$\begin{aligned} \int_a^b f(x) dx &= [F(c_1) - F(a)] + [F(c_2) - F(c_1)] + \dots + [F(b) - F(c_m)] = \\ &= F(b) - F(a), \end{aligned}$$

или, что то же,

$$\int_a^b f(x) dx = F(x) \Big|_a^b.$$

Примечание. Если функция $F(x)$, являющаяся первообразной для функции $f(x)$ на промежутках $[a, c_1)$, (c_1, c_2) , $\dots$, $(c_m, b]$, на отрезке $[a, b]$ не ограничена, то легко видеть, что интеграл $\int_a^b f(x) dx$ не существует.

Формальное применение формулы Ньютона — Лейбница в этом случае может привести к абсурдным результатам.

Например, интеграл $\int_{-1}^1 \frac{dx}{x^2}$ расходится, так как расходится интеграл

$$\int_0^1 \frac{dx}{x^2}.$$

Формальное применение формулы Ньютона — Лейбница в данном примере дает:

$$\int_{-1}^1 \frac{dx}{x^2} = -\frac{1}{x} \Big|_{-1}^1 = -1 - 1 = -2.$$

Это кажущееся противоречие объясняется тем, что первообразная $F(x) = -\frac{1}{x}$ функции $f(x) = \frac{1}{x^2}$ не ограничена вблизи точки $x=0$, принадлежащей отрезку интегрирования.

ГЛАВА VII

ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Раздел 1

ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ. ОСНОВНЫЕ ПОНЯТИЯ

§ 1. Функции нескольких переменных

До сих пор мы занимались изучением функции одной переменной, т. е. изучением переменной, значения которой зависят от значений одной независимой переменной.

На практике часто приходится иметь дело с величинами, численные значения которых зависят от значений нескольких изменяющихся независимо друг от друга физических и геометрических величин. Изучение таких величин приводит к понятию функции нескольких переменных.

Начнем с простейшего случая, когда таких переменных две.

Переменная z называется *функцией двух переменных* x и y , если каждой паре чисел (x, y) из некоторого множества по определенному правилу или закону ставится в соответствие одно или несколько значений переменной z .

При этом переменные x и y называются *независимыми переменными*, или *аргументами*, а переменная z — *зависимой переменной*, или *функцией*.

Множество пар чисел (x, y) , для которых определена функция z , называется *областью определения* этой функции.

Множество значений функции z называется *областью изменения* этой функции.

В математическом анализе обычно рассматриваются действительные функции двух действительных переменных, т. е. функции, принимающие действительные значения на множестве пар действительных чисел¹⁾.

Если каждой паре чисел (x, y) из области определения функции соответствует одно значение z , то функция называется *однозначной*, в противном случае — *многозначной*. Так же как и для функций одного переменного, если не оговорено противное, мы будем предполагать, что рассматриваемые функции двух переменных однозначны.

Тот факт, что переменная z является функцией переменных x и y , обычно записывают в виде:

$$z = f(x, y), \quad z = \varphi(x, y), \quad z = F(x, y) \quad \text{и т. д.}$$

¹⁾ Аналогичное замечание следует иметь в виду ниже, при изучении функций большего числа переменных.

Буквы f , φ , F и другие использованы здесь для обозначения закона соответствия между независимыми переменными x и y и зависимым переменным z ¹⁾. Частное значение функции $z=f(x, y)$ при $x=x_0$, $y=y_0$ обычно записывают в виде: $f(x_0, y_0)$ или $z|_{x=x_0, y=y_0}$.

Сам закон соответствия может быть задан произвольным образом (аналитически, т. е. с помощью формул, выражающих z через x и y , таблично, графически, словесным образом и пр.).

Математический анализ изучает преимущественно аналитически заданные функции.

Рассмотрим несколько примеров таких функций.

1. Если x и y могут принимать любые числовые значения, то переменные $z=x^2+y^2$, $z=1-x-y$, $z=\ln(1+x^4+y^4)$ и т. д. представляют собой аналитически заданные функции от x и y . Область определения каждой из них — множество всевозможных пар чисел (x, y) .

2. Объем конуса V есть функция его высоты H и радиуса основания R . Ее аналитическое выражение: $V=\frac{1}{3}\pi R^2 H$. Область определения — всевозможные пары чисел (R, H) , где $R>0$, $H>0$.

Естественной областью определения аналитически заданной функции $z=f(x, y)$ называется совокупность всех пар чисел (x, y) , которым соответствуют действительные значения функции. Так, например, для функции $z=\ln(x^2+y^2-1)$ естественная область определения состоит из всех пар чисел (x, y) , для которых $x^2+y^2-1>0$, т. е. $x^2+y^2>1$, а для функции $z=\sqrt{1-x^2}+\sqrt{4-y^2}$ естественная область определения состоит из всех пар чисел, для которых $1-x^2\geq 0$, $4-y^2\geq 0$, т. е. $-1\leq x\leq 1$, $-2\leq y\leq 2$.

Рассмотренная в примере 2 функция $V=\frac{1}{3}\pi R^2 H$ имеет естественной областью определения множество всевозможных пар чисел (R, H) .

В дальнейшем, если дополнительные ограничения на изменение независимых переменных постановкой задачи (как это было в примере 2) не накладываются, под областью определения аналитически заданной функции будем подразумевать ее естественную область определения.

Выберем на плоскости прямоугольную систему координат XOY и будем изображать пары чисел (x, y) точками плоскости с координатами x, y . Тогда область определения функции $f(x, y)$ будет изображаться некоторым множеством точек M плоскости, в связи с чем функцию двух переменных часто называют функцией точки M плоскости и обозначают $z=f(M)$, а ее область определения отождествляют с множеством изображающих ее точек.

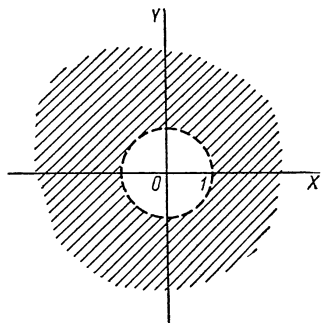


Рис. 283.

Так, для функции $z=\ln(x^2+y^2-1)$ область определения изображается множеством точек плоскости XOY , лежащих вне круга K с центром $(0, 0)$ радиуса 1, поэтому мы будем говорить: функция $z=\ln(x^2+y^2-1)$ определена вне круга K (см. рис. 283).

У функции $z=\sqrt{1-x^2}+\sqrt{4-y^2}$ область определения изображается множеством точек прямоугольника со сторонами $x=\pm 1$, $y=\pm 2$, поэтому мы будем говорить: функция $z=\sqrt{1-x^2}+\sqrt{4-y^2}$ определена в прямоугольнике со сторонами $x=\pm 1$, $y=\pm 2$ (рис. 284).

¹⁾ Иногда и закон соответствия и результат соответствия обозначают одной буквой, т. е. функцию записывают в виде:

$$z=z(x, y).$$

Как известно, функция одной переменной может быть изображена некоторой кривой на плоскости, если рассматривать значения ее аргумента как абсциссы, а значения функции как ординаты точек кривой.

Подобным же образом функция двух переменных $z = f(x, y)$ может быть геометрически изображена в виде поверхности.

Рассмотрим в пространстве прямоугольную систему координат XYZ . Область определения функции $z = f(x, y)$ изображается некоторым мно-

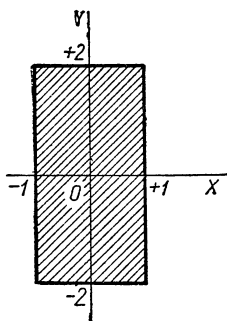


Рис. 284.

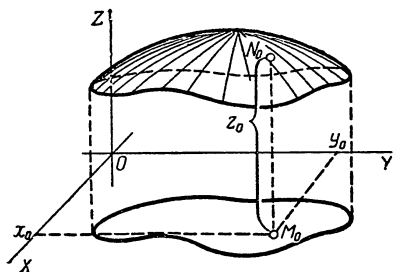


Рис. 285.

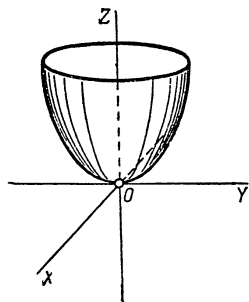


Рис. 286.

жеством точек плоскости HOY . Каждой точке $M_0(x_0, y_0)$ этого множества поставим в соответствие точку пространства $N_0(x_0, y_0, z_0)$, аппликата которой равна значению функции в точке M_0 : $z_0 = f(x_0, y_0)$. Совокупность всех таких точек представляет собой некоторую поверхность. Ее естественно принять за графическое изображение функции $z = f(x, y)$ (рис. 285). Итак, *графиком* функции $z = f(x, y)$ в пространстве XYZ является поверхность, представляющая собой геометрическое место точек $(x, y, f(x, y))$, когда точка (x, y) пробегает область определения функции.

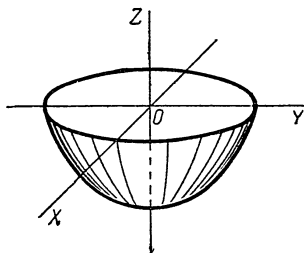


Рис. 287.

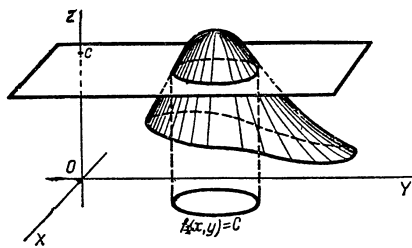


Рис. 288.

Например, для функции $z = x^2 + y^2$ область определения — вся плоскость, а изображающая ее поверхность — параболоид вращения (рис. 286); функция $z = -\sqrt{1 - x^2 - y^2}$ имеет областью определения круг $x^2 + y^2 \leq 1$ и изображается нижней полусферой с центром $O(0, 0, 0)$ радиуса $R = 1$ (рис. 287).

Познакомимся еще с одним способом геометрической иллюстрации функций двух переменных. Будем называть *линией уровня* функции $z = f(x, y)$ геометрическое место точек (x, y) плоскости, в которых функция принимает одно и то же значение C .

Линию уровня можно построить, спроектировав на плоскость HOY множество точек пространства XYZ , лежащих в пересечении поверхности, изображающей функцию $z = f(x, y)$, и плоскости $z = C$ (рис. 288).

Уравнение линии уровня имеет вид: $f(x, y) = C$. Изменяя C , мы будем получать различные линии уровня для данной функции.

Если положить $C = C_1, C_2, \dots, C_n, \dots$, выбрав эти числа в арифметической прогрессии с разностью h , то мы получим ряд линий уровня, по взаимному расположению которых можно судить о характере изменения функции (рис. 289). В частности, там, где линии гуще, функция изменяется быстрее (поверхность, изображающая функцию, идет круче), а там, где линии уровня располагаются реже, функция изменяется медленнее (соответствующая поверхность будет более полой). Кроме того, отметки на линиях уровня дают непосредственно значения функции в точках этих линий. Выбирая h достаточно малым, можно таким образом получить довольно точное представление о поведении функции.

Рассмотрим несколько примеров.

1. Функция $z = x^2 + y^2$ имеет линиями уровня окружности

$$x^2 + y^2 = C, \quad 0 \leq C < +\infty.$$

Полагая, например, $C = 0, 1, 2, 3, \dots$, получаем соответствующие линии уровня, сгущающиеся с ростом C . При $C = 0$ окружность вырождается в точку $(0, 0)$ (рис. 290).

Для функции $z = x + y$ линиями уровня будут прямые $x + y = C, -\infty < C < +\infty$. Полагая $C = 0, \pm 1, \pm 2, \dots$, получаем ряд линий уровня (рис. 291).

Линии уровня часто используются при составлении географических карт (линии уровня — линии, в которых высота точек земной поверхности над уровнем моря одинакова), при составлении метеорологических карт (линии уровня — линии одинаковых температур (изотермы), линии равного давления (изобары)) и т. д.

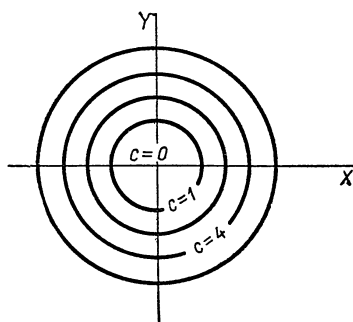


Рис. 290.

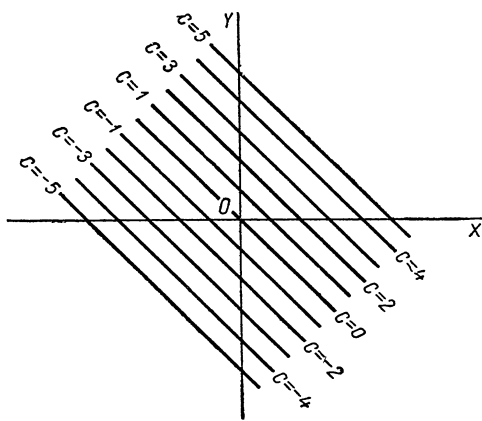


Рис. 291.

Понятие функции трех переменных дается аналогично случаю двух переменных.

Переменная u называется *функцией трех переменных* x, y, z^1 , если каждой тройке чисел (x, y, z) из некоторого множества по определенному правилу или закону ставится в соответствие одно или несколько значений переменной u . При этом x, y и z называются *независимыми переменными*, или *аргументами*, а u — *зависимой переменной*, или *функцией*.

¹⁾ Обозначается $u = f(x, y, z)$, $u = \varphi(x, y, z)$, $u = F(x, y, z)$, $u = u(x, y, z)$ и т. д.

Областью определения такой функции называют множество всех рассматриваемых троек чисел. Если функция задана аналитически, с помощью формул, под ее *естественной областью определения* подразумевают совокупность всех тех троек чисел (x, y, z) , для которых функция принимает действительные значения.

Изображая тройки чисел (x, y, z) точками пространства XYZ , мы можем рассматривать функцию трех переменных $u = f(x, y, z)$ как функцию точки $M(x, y, z)$ пространства, а область определения функции трех переменных — как некоторое множество точек пространства.

Так, например, функция $u = \frac{1}{\sqrt{4 - x^2 - y^2 - z^2}}$ определена для тех точек пространства, координаты которых удовлетворяют неравенству $4 - x^2 - y^2 - z^2 > 0$, т. е. определена внутри сферы $x^2 + y^2 + z^2 = 4$, а функция $u = x + y + z$ определена во всем пространстве.

Изобразить функцию трех переменных с помощью графика в трехмерном пространстве нельзя. Для наглядного изучения функций трех переменных используются так называемые поверхности уровня функции.

Поверхностью уровня функции $u = f(x, y, z)$ называют геометрическое место точек пространства, в которых функция принимает одно и то же значение C . Уравнение поверхности уровня: $f(x, y, z) = C$. Изменяя C , получаем различные поверхности уровня. По их взаимному расположению можно судить о характере поведения функции.

Определения функций двух и трех переменных легко переносятся на случай любого числа переменных. А именно, переменная u называется *функцией n переменных* $x, y, z, \dots, \omega$, если каждой системе из n чисел $(x, y, z, \dots, \omega)$ из некоторого множества по определенному правилу или закону ставится в соответствие одно или несколько значений переменной u . При этом переменные $x, y, z, \dots, \omega$ называются *независимыми переменными*, или *аргументами*, а переменная u — *зависимой переменной*, или *функцией*. Множество рассматриваемых систем из n чисел называется *областью определения* функции. Обозначаются функции n переменных в виде: $u = f(x, y, z, \dots, \omega)$, $u = F(x, y, z, \dots, \omega)$, $u = u(x, y, z, \dots, \omega)$ и т. д.

Так же как для $n = 2$ и $n = 3$, функции любого числа n переменных удобно рассматривать как функции точки n -мерного пространства. Будем называть *точкой M n -мерного пространства* всякую систему из n чисел $(x, y, z, \dots, \omega)$. Сами числа $x, y, z, \dots, \omega$ называются при этом *координатами* точки M . Совокупность всех таких точек составляет *n -мерное пространство*.

Расстояние между любыми двумя точками $M_1(x_1, y_1, z_1, \dots, \omega_1)$ и $M_2(x_2, y_2, z_2, \dots, \omega_2)$ этого пространства определяется по формуле

$$M_1 M_2 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 + \dots + (\omega_1 - \omega_2)^2}.$$

При $n = 1, 2, 3$ это «расстояние» совпадает с обычным. В n -мерном пространстве можно ввести понятия «прямых», кривых и пр., так чтобы при $n = 2, 3$, как частный случай, получались обычные определения.

Например, *прямой*, проходящей через точки $M_1(x_1, y_1, z_1, \dots, \omega_1)$ и $M_2(x_2, y_2, z_2, \dots, \omega_2)$, называют геометрическое место точек $(x, y, z, \dots, \omega)$, координаты которых удовлетворяют уравнениям:

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1} = \dots = \frac{\omega - \omega_1}{\omega_2 - \omega_1}.$$

Обозначая эти отношения через t , получаем уравнения той же прямой в параметрической форме:

$$\begin{aligned} x &= x_1 + t(x_2 - x_1), & y &= y_1 + t(y_2 - y_1), \\ z &= z_1 + t(z_2 - z_1), & \dots, & \omega = \omega_1 + t(\omega_2 - \omega_1). \end{aligned}$$

Непрерывной кривой в n -мерном пространстве называется любое геометрическое место точек $M(x, y, z, \dots, \omega)$, координаты которых удовлетворяют уравнениям:

$$x = x(t), \quad y = y(t), \quad z = z(t), \quad \dots, \quad \omega = \omega(t),$$

где $x(t), y(t), z(t), \dots, \omega(t)$ — непрерывные функции параметра t , изменяющегося в некоторых пределах.

Мы будем в дальнейшем подробно рассматривать лишь функции двух переменных, имея в виду, что перенос полученных определений и результатов на функции трех и более переменных представляет собой, как правило, лишь технические трудности.

§ 2. Понятие предела для функции двух и более переменных

Рассмотрим последовательность точек $M_1(x_1, y_1), M_2(x_2, y_2), \dots, M_n(x_n, y_n), \dots$ плоскости $ХОУ$. Будем говорить, что эта последовательность точек *сходится* к точке $M_0(x_0, y_0)$, если расстояние $M_n M_0 = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2}$ стремится к нулю при $n \rightarrow \infty$ ¹⁾.

Совокупность точек плоскости, находящихся от точки M_0 на расстоянии, меньшем δ , т. е. внутренность круга с центром M_0 радиуса δ , называется δ -окрестностью точки M_0 .

Таким образом, последовательность точек $M_1, M_2, \dots, M_n, \dots$ плоскости называется *сходящейся* к точке M_0 , если в любой окрестности точки M_0 лежат все точки последовательности, начиная с некоторого номера.

Обобщая данное ранее определение предела функции одного переменного с помощью предела числовой последовательности, дадим определение предела функции двух переменных.

Пусть функция $z = f(x, y)$ задана в некоторой окрестности точки M_0 , кроме, быть может, самой этой точки.

Если для любой последовательности точек $M_1(x_1, y_1), M_2(x_2, y_2), \dots, M_n(x_n, y_n), \dots$ этой окрестности²⁾, сходящейся к точке $M_0(x_0, y_0)$, соответствующая последовательность значений функции $f(x_1, y_1), f(x_2, y_2), \dots, f(x_n, y_n), \dots$ имеет пределом одно и то же число A , то это число A называют *пределом функции* $f(x, y)$ при $x \rightarrow x_0, y \rightarrow y_0$ и пишут:

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y) = A \text{ или } f(x, y) \rightarrow A \text{ при } x \rightarrow x_0, y \rightarrow y_0$$

Так, например, функция $f(x, y) = x^2 + y^2$ определена во всей плоскости. Рассмотрим точку $M_0(1, 2)$. Для любой последовательности точек $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), \dots$, сходящейся к этой точке M_0 , имеем:

$$\lim_{n \rightarrow \infty} (x_n^2 + y_n^2) = 1^2 + 2^2 = 5,$$

следовательно,

$$\lim_{\substack{x \rightarrow 1 \\ y \rightarrow 2}} (x^2 + y^2) = 5.$$

Геометрически тот факт, что число A является пределом функции $f(x, y)$ при $x \rightarrow x_0, y \rightarrow y_0$, означает, что, какова бы ни была последовательность точек плоскости $M_1, M_2, \dots, M_n, \dots$, неограниченно приближающихся к точке M_0 , последовательность аппликат соответствующих им точек поверхности, изображающей функцию $z = f(x, y)$, имеет пределом число A . При этом в самой точке M_0 функция может быть не определена, а может быть определена и иметь любое значение, как равное A , так и отличное от него. Предел функции $z = f(x, y)$ при $x \rightarrow x_0, y \rightarrow y_0$ определяется поведением функции вблизи точки (x_0, y_0) и не зависит от значения функции в самой этой точке.

В рассмотренном выше примере мы имели:

$$\begin{aligned} f(x, y) &= x^2 + y^2, \\ f(1, 2) &= 1^2 + 2^2 = 5 \text{ и } \lim_{\substack{x \rightarrow 1 \\ y \rightarrow 2}} f(x, y) = 5; \end{aligned}$$

¹⁾ Или, что то же, если $x_n \rightarrow x_0, y_n \rightarrow y_0$ при $n \rightarrow \infty$.

²⁾ Точки $M_1, M_2, \dots, M_n, \dots$ предполагаются отличными от M_0 .

для функции же $\varphi(x, y)$, определяемой равенствами

$$\varphi(x, y) = \begin{cases} 7 & \text{в точке } (1, 2), \\ f(x, y) & \text{во всех остальных точках,} \end{cases}$$

$\varphi(1, 2) = 7$, тогда как предел функции $\varphi(x, y)$ при $x \rightarrow 1, y \rightarrow 2$ равен пределу функции $f(x, y)$.

Приведем теперь пример функции, не имеющей предела при $x \rightarrow x_0, y \rightarrow y_0$. Функция $z = \frac{x^2 - y^2}{x^2 + y^2}$, определенная всюду, кроме начала координат, при $x \rightarrow 0, y \rightarrow 0$ предела не имеет, так как, выбирая две сходящихся к точке $O(0, 0)$ последовательности точек $(x_1, 0), (x_2, 0), \dots, (x_n, 0), \dots$ и $(0, y_1), (0, y_2), \dots, (0, y_n), \dots$, получим:

$$\lim_{n \rightarrow \infty} \frac{x_n^2 - 0}{x_n^2 + 0} = 1, \quad \lim_{n \rightarrow \infty} \frac{0 - y_n^2}{0 + y_n^2} = -1.$$

Так же как для функции одной переменной, можно показать, что данное выше определение предела функции двух переменных эквивалентно следующему определению предела: число A называется *пределом* функции $f(x, y)$ при $x \rightarrow x_0, y \rightarrow y_0$, если для любого положительного числа ε существует число $\delta, \delta > 0$, такое, что для всех точек (x, y) , отличных от точки (x_0, y_0) и удовлетворяющих условию $\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$, выполнено неравенство

$$|f(x, y) - A| < \varepsilon.$$

Геометрически это значит, что, каково бы ни было число $\varepsilon > 0$, найдется столь малая δ -окрестность точки $M_0(x_0, y_0)$, что во всех ее точках $M(x, y)$, отличных от M_0 , аппликаты соответствующих точек поверхности, изображающей функцию $z = f(x, y)$, отличаются от числа A по абсолютной величине меньше, чем на ε .

Величина δ зависит от ε и при ε , стремящемся к нулю, δ , вообще говоря, также стремится к нулю.

Будем называть функцию $f(x, y)$ *бесконечно малой* при $x \rightarrow x_0, y \rightarrow y_0$, если

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y) = 0.$$

Пользуясь любым из определений предела функции двух переменных, можно вывести основные свойства бесконечно малых функций, дать понятие порядка, эквивалентности бесконечно малых; доказать теорему о том, что разность между функцией, имеющей предел, и ее пределом есть бесконечно малая функция; доказать основные теоремы об арифметических операциях над пределами. Доказательства этих теорем аналогичны соответствующим доказательствам для функций одной переменной.

Не останавливаясь на понятии бесконечного предела функции двух переменных (его с успехом даст сам читатель по аналогии со случаем одной переменной), отметим, что в дальнейшем, говоря о пределах, мы будем иметь в виду *конечный* предел.

Данные выше определения без труда обобщаются на функции трех и более переменных, теряется лишь их геометрическая наглядность.

Так, например, δ -окрестностью точки $(x_0, y_0, z_0, \dots, w_0)$ n -мерного пространства называется совокупность точек $(x, y, z, \dots, w)$, находящихся от точки $(x_0, y_0, z_0, \dots, w_0)$ на расстоянии, меньшем δ .

В случае $n = 3$ δ -окрестность точки (x_0, y_0, z_0) представляет собой внутренность шара радиуса δ с центром в этой точке.

Для функции $f(x, y, z, \dots, w)$, определенной в некоторой окрестности точки $(x_0, y_0, z_0, \dots, w_0)$ (кроме, быть может, этой точки), можно дать, например, следующее определение предела. Число A называется *пределом* функции $f(x, y, z, \dots, w)$ при $x \rightarrow x_0, y \rightarrow y_0, z \rightarrow z_0, \dots, w \rightarrow w_0$, если для любого числа $\varepsilon > 0$ существует число

$\delta > 0$, такое, что для всех точек $(x, y, z, \dots, w)$ указанной окрестности, отличных от точки $(x_0, y_0, z_0, \dots, w_0)$ и удовлетворяющих условию

$$\sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 + \dots + (w-w_0)^2} < \delta,$$

выполнено неравенство

$$|f(x, y, z, \dots, w) - A| < \varepsilon.$$

§ 3. Непрерывность функций нескольких переменных

Функция $z = f(x, y)$, определенная в некоторой окрестности точки (x_0, y_0) , называется *непрерывной в этой точке*, если

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y) = f(x_0, y_0).$$

Геометрически это означает, что при приближении точки $M(x, y)$ по любой последовательности точек к точке $M_0(x_0, y_0)$ аппликаты соответствующих точек поверхности, изображающей функцию $z = f(x, y)$, стремятся к аппликате поверхности в точке M_0 .

В дальнейшем нам часто придется пользоваться следующим эквивалентным данному выше определением непрерывности функции: функция $f(x, y)$ называется *непрерывной в точке (x_0, y_0)* ¹⁾, если для любого положительного числа ε существует число $\delta > 0$, такое, что для всех точек (x, y) , удовлетворяющих условию $\sqrt{(x-x_0)^2 + (y-y_0)^2} < \delta$, выполнено неравенство

$$|f(x, y) - f(x_0, y_0)| < \varepsilon.$$

Пользуясь определением непрерывности и теоремами о пределах, можно доказать, что для функций двух переменных сумма и произведение двух непрерывных функций есть непрерывная функция; частное двух непрерывных функций есть непрерывная функция в точках, в которых знаменатель отличен от нуля; сложная функция²⁾, составленная из непрерывных функций, есть непрерывная функция и т. д.

Например, функции $f(x, y) = x^2 + y^2$, $\varphi(x, y) = e^{x+y}$, $F(x, y) = \sin(x^2 y^2 + x - y + 1)$ непрерывны всюду, функция $R(x, y) = \frac{x^2 + y^2 - 1}{x - y}$ непрерывна всюду, кроме точек прямой $x = y$, и т. д.

Если функция $f(x, y)$ определена в окрестности точки $M_0(x_0, y_0)$ (кроме, может быть, самой точки M_0) и не является непрерывной в этой точке, будем называть ее *разрывной* в точке M_0 .

Например, функция $f(x, y) = \frac{1}{x^2 + y^2}$ разрывна в точке $(0, 0)$, так как она определена всюду, кроме этой точки; функция

$$f(x, y) = \begin{cases} \frac{x^2 - y^2}{x^2 + y^2} & \text{— всюду, кроме точки } (0, 0), \\ 1 & \text{— в точке } (0, 0), \end{cases}$$

будет разрывна в точке $(0, 0)$, так как не существует предела этой функции при $x \rightarrow 0, y \rightarrow 0$; функция

$$\varphi(x, y) = \begin{cases} x^2 + y^2 & \text{— всюду, кроме точки } (1, 2), \\ 7 & \text{— в точке } (1, 2), \end{cases}$$

в точке $(1, 2)$ разрывна, так как

$$\lim_{\substack{x \rightarrow 1 \\ y \rightarrow 2}} \varphi(x, y) = 5, \text{ а } \varphi(1, 2) = 7.$$

В гл. III, § 22 мы рассматривали свойства функций, непрерывных на отрезках (теоремы об ограниченности функции, о достижении наи-

¹⁾ Функция снова предполагается заданной в некоторой окрестности точки $M_0(x_0, y_0)$.

²⁾ Подробнее о сложных функциях см. в § 8.

больших и наименьших значений и т. д.). Аналогичные свойства имеют место для непрерывных функций двух переменных, если их рассматривать в так называемых замкнутых ограниченных областях, являющихся двумерным аналогом отрезка. Для того чтобы сформулировать эти свойства, введем ряд определений для множеств точек плоскости XOY .

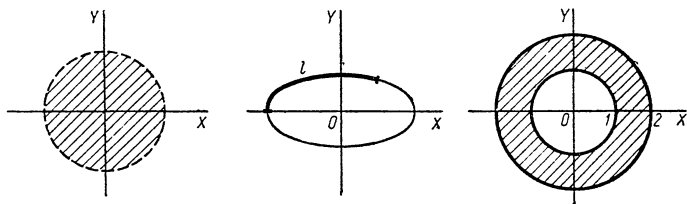


Рис. 292.

Множество точек плоскости называется *связным*, если любые две точки этого множества можно соединить непрерывной кривой, состоящей из точек того же множества.

Например, связным будет множество точек внутри круга, множество точек любой дуги l эллипса, множество точек кольца $1 \leq x^2 + y^2 \leq 4$ (рис. 292) и т. д.

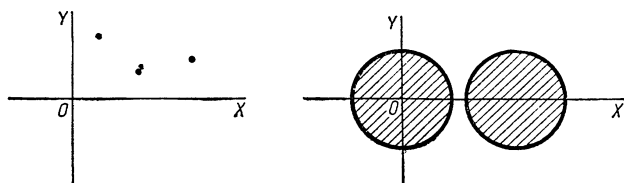


Рис. 293.

Множество точек, состоящее из трех отдельных точек, не будет связным. Несвязным будет также множество точек, принадлежащих двум кругам: $x^2 + y^2 \leq 1$ и $(x-2,3)^2 + y^2 \leq 1$ (рис. 293) и т. д.

Точка M называется *внутренней* точкой некоторого множества, если существует окрестность этой точки, состоящая из точек данного множества.

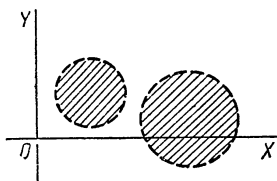


Рис. 294.

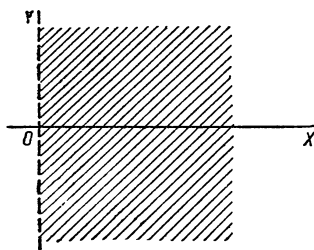


Рис. 295.

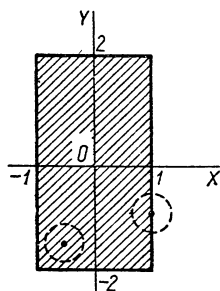


Рис. 296.

Множество, каждая точка которого — внутренняя, называется *открытым* множеством.

Открытыми будут, например, множества всех точек, лежащих внутри круга, внутри двух непересекающихся кругов (рис. 294). Открытым будет множество всех точек плоскости, лежащих вправо от оси OY (рис. 295).

Множество точек прямоугольника $-1 \leq x \leq 1$, $-2 \leq y \leq 2$ открытым не будет, так как принадлежащие ему точки прямых $x = \pm 1$, $y = \pm 2$ не являются внутренними (рис. 296).

Связное открытое множество точек плоскости называется *областью*. Примеры областей: множества точек, лежащих внутри круга, прямоугольника, множество точек кольца $1 < x^2 + y^2 < 4$, множество точек (x, y) плоскости, для которых $x > 0, y > 0$ (рис. 297).

Точка M называется *граничной* точкой области, если в любой ее окрестности есть точки как принадлежащие, так и не принадлежащие этой области. Совокупность всех граничных точек области называется ее *границей*. Так, например, для области, состоящей из точек, лежащих внутри эллипса, границей будет сам эллипс.

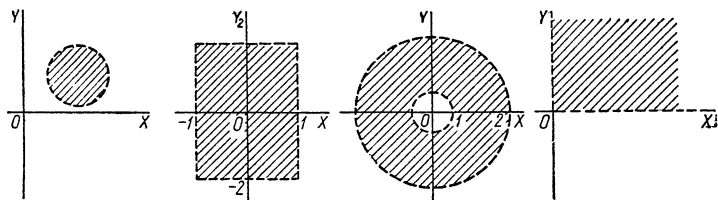


Рис. 297.

Область с присоединенной к ней границей называется *замкнутой* областью. Замкнутыми областями будут, например, круг $x^2 + y^2 \leq 1$; прямоугольник $-1 \leq x \leq 1, -2 \leq y \leq 2$; кольцо $1 \leq x^2 + y^2 \leq 4$; бесконечный сектор, состоящий из точек (x, y) плоскости, для которых $x \geq 0, y \geq 0$ и т. д.

Область называется *ограниченной*, если она целиком лежит внутри некоторого круга с центром в начале координат.

В указанных выше примерах областей все области, кроме последней, ограниченные (рис. 297).

Пусть функция $f(x, y)$ определена в некоторой области D (замкнутой или незамкнутой). Будем называть эту функцию *непрерывной в области D* , если она непрерывна в каждой ее точке. При этом непрерывность в любой граничной точке области D определяется следующим образом: функция $f(x, y)$ *непрерывна в граничной точке $M_0(x_0, y_0)$* , если для любого положительного числа ε существует число $\delta > 0$, такое, что для всех точек $M(x, y)$ области D , удовлетворяющих условию $\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$, выполнено неравенство $|f(x, y) - f(x_0, y_0)| < \varepsilon$.

Функции, непрерывные в замкнутых ограниченных областях, обладают следующими свойствами.

1. Функция $f(x, y)$, непрерывная в замкнутой ограниченной области, ограничена в этой области, т. е. существует число K , такое, что для всех точек области $|f(x, y)| \leq K$.

2. Функция $f(x, y)$, непрерывная в замкнутой ограниченной области, принимает в этой области свое наибольшее и наименьшее значения. Это значит, что существуют числа m и M , такие, что для всех точек (x, y) области

$$m \leq f(x, y) \leq M,$$

причем в области имеются точки (x_1, y_1) и (x_2, y_2) , такие, что $f(x_1, y_1) = m$, $f(x_2, y_2) = M$. Здесь m — наименьшее, а M — наибольшее значения функции $f(x, y)$ в данной области.

3. Функция $f(x, y)$, непрерывная в замкнутой ограниченной области, между любыми двумя своими значениями принимает все промежуточные, т. е. если $A < C < B$, где A и B — какие-то значения функции $f(x, y)$ в данной области, то в этой области существует точка (x_0, y_0) , в которой $f(x_0, y_0) = C$.

Отсюда, в частности, следует, что если (x_1, y_1) и (x_2, y_2) — точки данной области, причем $f(x_1, y_1) < 0$, а $f(x_2, y_2) > 0$, то в области существует точка (x_0, y_0) , в которой $f(x_0, y_0) = 0$.

4. Если функция $f(x, y)$ непрерывна в замкнутой ограниченной области, то она *равномерно непрерывна* в этой области, т. е. для любого положительного числа ε существует число $\delta > 0$, такое, что для любых двух точек (x_1, y_1) и (x_2, y_2) области, находящихся на расстоянии, меньшем δ , выполнено неравенство

$$|f(x_1, y_1) - f(x_2, y_2)| < \varepsilon.$$

Доказательства перечисленных в пп. 1—4 теорем мы в нашем курсе не рассматриваем. Заметим только, что для функций, непрерывных в незамкнутых или неограниченных областях, указанные в пп. 1, 2, 4 свойства могут не иметь места.

Понятия непрерывности функции в точке, в области и перечисленные свойства непрерывных функций двух переменных легко обобщаются на функции трех и более переменных.

Раздел 2

ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИАЛЫ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 4. Частные производные

Пусть в некоторой (открытой) области задана функция двух переменных $z = f(x, y)$. Возьмем произвольную точку $M(x, y)$ этой области и дадим x приращение Δx , остав-
ляя значение y неизменным.

При этом функция $f(x, y)$ получит приращение

$$\Delta_x z = f(x + \Delta x, y) - f(x, y).$$

Оно называется *частным приращением* этой функции по x (рис. 298).

Отношение

$$\frac{\Delta_x z}{\Delta x} = \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

дает среднюю скорость изменения функции $z = f(x, y)$ по переменной x на участке от точки $M(x, y)$ до точки $M'(x + \Delta x, y)$.

Предел этого отношения при $\Delta x \rightarrow 0$, если он существует и конечен, называется *частной производной функции $z = f(x, y)$ по переменной x в точке (x, y)* . Этот предел характеризует скорость изменения данной функции по x в самой точке M . Частную производную по x от функции $z = f(x, y)$ обозначают следующими символами:

$$\frac{\partial z}{\partial x}, \quad z'_x, \quad \frac{\partial f(x, y)}{\partial x}, \quad f'_x(x, y).$$

$$\text{Таким образом, } \frac{\partial z}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta_x z}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}.$$

Аналогично, считая x постоянной и давая y приращение Δy , мы получим *частное приращение* функции $z = f(x, y)$ по y :

$$\Delta_y z = f(x, y + \Delta y) - f(x, y) \quad (\text{рис. 299}).$$

Предел отношения частного приращения функции по y к приращению Δy при стремлении последнего к нулю (если он существует и конечен) называется *частной производной этой функции по y в точке (x, y)* .

Частная производная функции $z=f(x, y)$ по y обозначается обычно одним из следующих символов:

$$\frac{\partial z}{\partial y}, \quad z'_y, \quad \frac{df(x, y)}{\partial y}, \quad f'_y(x, y).$$

Таким образом,

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{\Delta_y z}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}.$$

Эта частная производная численно равна скорости изменения по y функции $z=f(x, y)$ в точке $M(x, y)$.

Значения частных производных $f'_x(x, y)$ и $f'_y(x, y)$, естественно, зависят от координат x, y рассматриваемой точки M , т. е. частные производные $f'_x(x, y)$ и $f'_y(x, y)$, в свою очередь, являются функциями двух переменных, определенными в заданной области или ее части (если не во всех точках этой области существуют частные производные).

Вычисление частных производных по x (по y) от конкретных функций производится по известным для функций одной переменной правилам, так как частная производная функции $z=f(x, y)$ по x (по y) есть по определению обыкновенная производная функции $z=f(x, y)$, рассматриваемой как функции одной переменной x (соответственно y) при постоянном значении другой переменной.

Например, для функции $f(x, y) = x^2 + xy^2 + y^3$ в любой точке (x, y) имеем $f'_x(x, y) = 2x + y^2$, $f'_y(x, y) = 2xy + 3y^2$. В частности, $f'_x(1, 2) = 2 \cdot 1 + 2^2 = 6$, $f'_y(1, 2) = 2 \cdot 1 \cdot 2 + 3 \cdot 2^2 = 16$.

Частные производные функции двух переменных имеют простой геометрический смысл.

Рассмотрим в пространстве XYZ поверхность S , имеющую уравнение $z=f(x, y)$. Эта поверхность изображает функцию $f(x, y)$.

Дадим переменным x и y значения x_0, y_0 . По определению $f'_x(x_0, y_0)$ есть обыкновенная производная (по x) от функции одной переменной $f(x, y_0)$ при

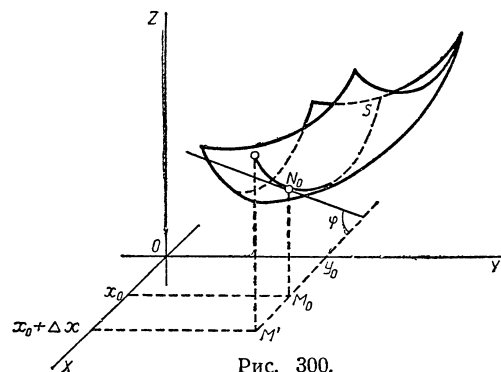


Рис. 300.

$x=x_0$. График этой функции мы получим в сечении поверхности S плоскостью $y=y_0$.

Так как производная функции одной переменной x в данной точке равна тангенсу угла наклона к оси OX касательной к графику функции в этой точке, заключаем, что $f'_x(x_0, y_0)$ есть тангенс угла φ , составленного с осью OX касательной, проведенной в точке $N_0(x_0, y_0, f(x_0, y_0))$ к сечению поверхности S плоскостью $y=y_0$ (рис. 300).

Аналогично $f'_y(x_0, y_0)$ есть тангенс угла θ , составляемого с осью OY касательной, проведенной в точке N_0 к сечению поверхности S плоскостью $x=x_0$ (рис. 301).

Для функций любого числа n переменных частные производные вводятся так же, как и для функций двух переменных. А именно, част-

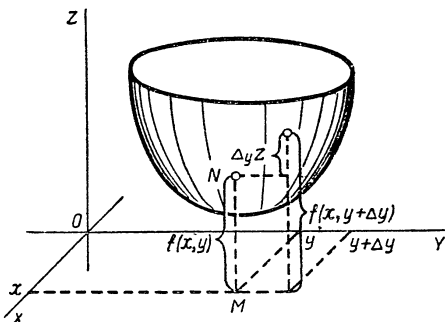


Рис. 299.

ная производная от функции $u = f(x, y, z, \dots, w)$ по любому из независимых переменных в точке $(x, y, z, \dots, w)$ есть предел отношения частного приращения функции в этой точке к приращению соответствующей независимой переменной при стремлении последнего к нулю (если этот предел существует и конечен). Обозначения аналогичны приведенным выше.

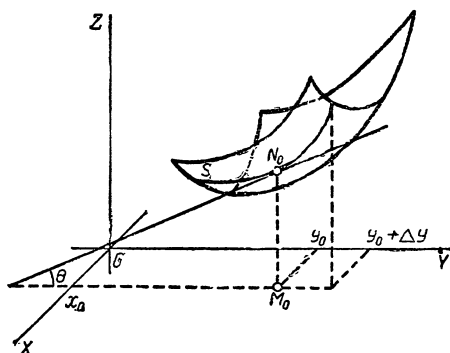


Рис. 301.

Например, если $u = f(x, y, z)$, то

$$\frac{\partial u}{\partial z} = f_z(x, y, z) = \lim_{\Delta z \rightarrow 0} \frac{\Delta_z u}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{f(x, y, z + \Delta z) - f(x, y, z)}{\Delta z}.$$

§ 5. Дифференцируемые функции нескольких переменных

Рассмотрим снова функцию двух переменных $z = f(x, y)$ в некоторой (открытой) области. Пусть (x, y) — точка этой области. Перейдем из точки (x, y) в какую-нибудь точку $(x + \Delta x, y + \Delta y)$ той же области. При этом функция $z = f(x, y)$ получит приращение

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

В отличие от частных приращений $\Delta_x z$ и $\Delta_y z$ это приращение называется *полным приращением* функции $z = f(x, y)$ в точке (x, y) , соответствующим приращениям Δx и Δy независимых переменных. Геометрически Δz представляет собой разность аппликат точек поверхности $z = f(x, y)$, соответствующих точкам (x, y) и $(x + \Delta x, y + \Delta y)$ (рис. 302).

В зависимости от вида функции и величин Δx и Δy полное приращение функции может быть больше нуля, меньше или равно нулю.

Если функция $z = f(x, y)$ непрерывна в точке (x, y) , то

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} f(x + \Delta x, y + \Delta y) = f(x, y),$$

т. е. $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \Delta z = 0$, и, наоборот, если $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \Delta z = 0$, то, как легко видеть, функция $z = f(x, y)$ непрерывна в точке (x, y) .

Функция $z = f(x, y)$ называется *дифференцируемой в точке (x, y)* , если ее полное приращение в этой точке можно представить в виде:

$$\Delta z = A(x, y) \Delta x + B(x, y) \Delta y + \alpha_1 \Delta x + \alpha_2 \Delta y, \quad (1)$$

где α_1 и α_2 — бесконечно малые при $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$ (или, короче, при $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2} \rightarrow 0$ ¹⁾).

Соотношение (1) часто записывают в более сжатой форме:

$$\Delta z = A(x, y) \Delta x + B(x, y) \Delta y + \alpha \rho. \quad (1')$$

Здесь $\alpha = \frac{\alpha_1 \Delta x + \alpha_2 \Delta y}{\rho} \rightarrow 0$ при $\rho \rightarrow 0$.

¹⁾ Здесь α_1 и α_2 зависят от $x, y, \Delta x$ и Δy .

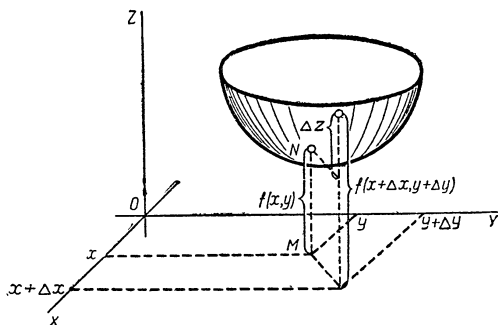


Рис. 302.

При этом слагаемое $A(x, y) \Delta x + B(x, y) \Delta y$, линейное относительно Δx и Δy , называется *главной частью* приращения, так как оставшееся слагаемое $\alpha_1 \Delta x + \alpha_2 \Delta y$, т. е. $\alpha\rho$, есть бесконечно малая более высокого порядка, чем Δx и Δy (или, короче, чем ρ).

Рассмотрим пример. Функция $z = x^2 y$ будет дифференцируемой в любой точке (x, y) , так как

$$\Delta z = (x + \Delta x)^2 (y + \Delta y) - x^2 y = 2xy \Delta x + x^2 \Delta y + 2x \Delta x \Delta y + y (\Delta x)^2 + (\Delta x)^2 \Delta y.$$

Здесь $2xy \Delta x + x^2 \Delta y$ — главная линейная относительно Δx и Δy часть полного приращения функции, а слагаемое $(2x \Delta x + y \Delta x) \Delta x + (\Delta x)^2 \Delta y$ есть бесконечно малая высшего порядка по сравнению с Δx и Δy .

Данное выше определение дифференцируемости для функции двух переменных является естественным обобщением определения дифференцируемости для функции одной переменной¹⁾.

В гл. IV было установлено, что если функция одной переменной дифференцируема в некоторой точке, то она непрерывна и имеет производную в этой точке. Последнее условие оказалось и достаточным, т. е. из существования производной функции одной переменной в данной точке следует дифференцируемость функции в этой точке. Выясним, как переносятся эти свойства на функции двух переменных.

Необходимые условия дифференцируемости

Теорема 1. Если функция $z = f(x, y)$ дифференцируема в точке (x, y) , то она непрерывна в этой точке.

Доказательство. Если функция $z = f(x, y)$ дифференцируема в точке (x, y) , то, как следует из соотношения (1), $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \Delta z = 0$, а это

значит, что функция $f(x, y)$ непрерывна в точке (x, y) .

Теорема 2. Если функция $z = f(x, y)$ дифференцируема в точке (x, y) , то она имеет в этой точке частные производные $f'_x(x, y)$ и $f'_y(x, y)$.

Доказательство. Так как функция $z = f(x, y)$ дифференцируема в точке (x, y) , то для любых Δx и Δy (не выводящих за пределы рассматриваемой окрестности) имеет место соотношение (1). Полагая в нем $\Delta y = 0$, имеем $\Delta_x z = A(x, y) \Delta x + \alpha_1 \Delta x$, где α_1 — бесконечно малая при $\Delta x \rightarrow 0$.

Деля на Δx и переходя к пределу при $\Delta x \rightarrow 0$, получаем:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta_x z}{\Delta x} = \lim_{\Delta x \rightarrow 0} [A(x, y) + \alpha_1] = A(x, y).$$

Следовательно, в точке (x, y) существует частная производная $f'_x(x, y)$ и $f'_x(x, y) = A(x, y)$.

Аналогично доказывается, что в точке (x, y) существует частная производная $f'_y(x, y)$ и что $f'_y(x, y) = B(x, y)$.

Обратные утверждения к теоремам 1 и 2 неверны. Из непрерывности функции двух переменных в точке, а также из существования ее частных производных в этой точке еще не следует дифференцируемости функции.

Действительно, функция $f(x, y) = x + y + \sqrt{|x| \cdot |y|}$ непрерывна в точке $(0, 0)$, так как $f(0, 0) = 0$ и $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} f(x, y) = 0$, и имеет в этой точке частные производные:

$$f'_x(0, 0) = \lim_{\Delta x \rightarrow 0} \frac{(0 + \Delta x) + 0 + \sqrt{|\Delta x| \cdot 0} - 0}{\Delta x} = 1,$$

$$f'_y(0, 0) = \lim_{\Delta y \rightarrow 0} \frac{0 + (0 + \Delta y) + \sqrt{0 \cdot |\Delta y|} - 0}{\Delta y} = 1.$$

¹⁾ Напомним, что функцию $y = f(x)$ мы называем дифференцируемой в точке x , если ее приращение можно представить в виде: $\Delta y = A(x) \Delta x + \alpha \Delta x$, где α — бесконечно малая при $\Delta x \rightarrow 0$.

Ее полное приращение в точке $(0, 0)$ равно:

$$\Delta z = \Delta x + \Delta y + \sqrt{|\Delta x| \cdot |\Delta y|}.$$

Если бы функция $f(x, y) = x + y + \sqrt{|x| \cdot |y|}$ была дифференцируема в точке $(0, 0)$, мы имели бы $\Delta z = 1 \cdot \Delta x + 1 \cdot \Delta y + \alpha \cdot \rho$, где α — бесконечно малая при $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2} \rightarrow 0$.

Но у нас $\alpha \cdot \rho = \sqrt{|\Delta x| \cdot |\Delta y|}$, т. е. $\alpha = \frac{\sqrt{|\Delta x| \cdot |\Delta y|}}{\sqrt{(\Delta x)^2 + (\Delta y)^2}}$. При любых $\Delta x = \Delta y$ имеем $\alpha = \frac{1}{\sqrt{2}}$, а значит, α не является бесконечно малой при $\Delta x \rightarrow 0, \Delta y \rightarrow 0$.

Таким образом, функция $f(x, y) = x + y + \sqrt{|x| \cdot |y|}$ в точке $(0, 0)$ непрерывна, имеет частные производные и, тем не менее, не является дифференцируемой в этой точке.

Для того чтобы функция двух переменных была дифференцируема в данной точке, на нее, в отличие от функции одной переменной, надо наложить более жесткие требования, чем существование частных производных в этой точке.

Достаточные условия дифференцируемости

Теорема. Если в некоторой окрестности точки (x, y) существуют частные производные $f'_x(x, y)$ и $f'_y(x, y)$ функции $z = f(x, y)$ и эти производные непрерывны в самой точке (x, y) , то функция $f(x, y)$ дифференцируема в этой точке.

Доказательство. Для любых Δx и Δy (не выводящих за пределы рассматриваемой окрестности) имеем:

$$\begin{aligned} \Delta z &= f(x + \Delta x, y + \Delta y) - f(x, y) = \\ &= [f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)] + [f(x, y + \Delta y) - f(x, y)]. \end{aligned}$$

Разность $f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)$ можно рассматривать как приращение функции одной переменной x : $f(x, y + \Delta y)$ при переходе от x к $x + \Delta x$. Так как на отрезке $[x, x + \Delta x]$ (или $[x + \Delta x, x]$, если $\Delta x < 0$) эта функция имеет производную, совпадающую с $f'_x(x, y + \Delta y)$, то по теореме Лагранжа получаем:

$$f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) = f'_x(x + \theta_1 \Delta x, y + \Delta y) \Delta x,$$

где $0 < \theta_1 < 1$.

Аналогично для второй разности получаем:

$$f(x, y + \Delta y) - f(x, y) = f'_y(x, y + \theta_2 \Delta y) \Delta y,$$

где $0 < \theta_2 < 1$.

Таким образом, $\Delta z = f'_x(x + \theta_1 \Delta x, y + \Delta y) \Delta x + f'_y(x, y + \theta_2 \Delta y) \Delta y$, где $0 < \theta_1 < 1, 0 < \theta_2 < 1$.

Функции $f'_x(x, y)$ и $f'_y(x, y)$ непрерывны в точке (x, y) , значит,

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} f'_x(x + \theta_1 \Delta x, y + \Delta y) = f'_x(x, y),$$

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} f'_y(x, y + \theta_2 \Delta y) = f'_y(x, y),$$

т. е.

$$f'_x(x + \theta_1 \Delta x, y + \Delta y) = f'_x(x, y) + \alpha_1, \quad f'_y(x, y + \theta_2 \Delta y) = f'_y(x, y) + \alpha_2,$$

где α_1 и α_2 — бесконечно малые при $\Delta x \rightarrow 0, \Delta y \rightarrow 0$.

Подставляя эти значения в формулу (2) для Δz , получаем:

$$\Delta z = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y + \alpha_1 \Delta x + \alpha_2 \Delta y,$$

где α_1 и α_2 — бесконечно малые при $\Delta x \rightarrow 0, \Delta y \rightarrow 0$, а это и значит, что функция $f(x, y)$ дифференцируема в точке (x, y) .

С помощью только что доказанной теоремы мы можем устанавливать дифференцируемость для широкого класса функций.

Так, например, функция $z = x^2 e^{x^2 y^3}$ в любой точке (x, y) дифференцируема, так как ее частные производные $\frac{\partial z}{\partial x} = (2x + 2x^3 y^3) e^{x^2 y^3}$ и $\frac{\partial z}{\partial y} = 3x^4 y^2 e^{x^2 y^3}$ всюду непрерывны; функция $z = \sqrt{x+y}$ дифференцируема в каждой точке полуплоскости $x+y > 0$, так как там существуют и непрерывны ее частные производные $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = \frac{1}{2\sqrt{x+y}}$, и т. д.

Понятие дифференцируемости для функций трех и большего числа переменных вводится совершенно аналогично рассмотренному выше случаю функции двух переменных.

Функция $u = f(x, y, z, \dots, w)$, определенная в некоторой окрестности точки $(x, y, z, \dots, w)$ n -мерного пространства, называется *дифференцируемой* в этой точке, если ее полное приращение в этой точке можно представить в виде:

$$\Delta u = A(x, y, z, \dots, w) \Delta x + B(x, y, z, \dots, w) \Delta y + C(x, y, z, \dots, w) \Delta z + \dots + M(x, y, z, \dots, w) \Delta w + \alpha_1 \Delta x + \alpha_2 \Delta y + \alpha_3 \Delta z + \dots + \alpha_n \Delta w,$$

где $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$ — бесконечно малые при $\Delta x \rightarrow 0, \Delta y \rightarrow 0, \Delta z \rightarrow 0, \dots, \Delta w \rightarrow 0$. Слагаемое $\alpha_1 \Delta x + \alpha_2 \Delta y + \alpha_3 \Delta z + \dots + \alpha_n \Delta w$ часто записывают в виде $\alpha \rho$, причем для дифференцируемой функции $\alpha \rightarrow 0$ при $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 + \dots + (\Delta w)^2} \rightarrow 0$.

Имеют место следующие необходимые (теорема 1') и достаточные (теорема 2') условия дифференцируемости.

Теорема 1'. Функция $f(x, y, z, \dots, w)$, дифференцируемая в точке $(x, y, z, \dots, w)$, непрерывна в этой точке и имеет в ней частные производные по всем независимым переменным.

Теорема 2'. Если функция $f(x, y, z, \dots, w)$ в некоторой окрестности точки $(x, y, z, \dots, w)$ имеет частные производные по переменным $x, y, z, \dots, w$ и если эти частные производные непрерывны в самой точке $(x, y, z, \dots, w)$, то функция $f(x, y, z, \dots, w)$ дифференцируема в этой точке.

Доказательства этих теорем не содержат ничего принципиально нового по сравнению со случаем функций двух переменных.

Функция (любого числа переменных), дифференцируемая в каждой точке некоторой области, называется *дифференцируемой* в этой области.

§ 6. Дифференциал

Если функция $z = f(x, y)$ дифференцируема в точке (x, y) , то, как было показано выше, ее полное приращение в этой точке можно представить в виде:

$$\Delta z = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y + \alpha_1 \Delta x + \alpha_2 \Delta y,$$

где α_1 и α_2 — бесконечно малые при $\Delta x \rightarrow 0, \Delta y \rightarrow 0$.

Главная, линейная относительно Δx и Δy часть приращения дифференцируемой функции $z = f(x, y)$ в точке (x, y) , т. е. $f'_x(x, y) \Delta x + f'_y(x, y) \Delta y$, называется ее *полным дифференциалом* в этой точке и обозначается $df(x, y)$ или dz .

Итак,

$$dz = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y.$$

Например, если $z = x^2 y$, то в любой точке (x, y) $dz = 2xy \Delta x + x^2 \Delta y$.

По определению разность между полным приращением и полным дифференциалом функции в данной точке есть бесконечно малая более высокого порядка, чем Δx и Δy (чем $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$):

$$\Delta z - dz = \alpha_1 \Delta x + \alpha_2 \Delta y \text{ или, что то же, } \Delta z - dz = \alpha \rho,$$

где α_1, α_2 и α — бесконечно малые при $\rho \rightarrow 0$.

Этим широко пользуются в приближенных вычислениях, заменяя приращение функции $z = f(x, y)$, которое может сложным образом зависеть от Δx и Δy , ее дифференциалом, зависящим от Δx и Δy линейно.

Пусть, например, нам известны значения дифференцируемой функции $z = f(x, y)$ и ее частных производных $f'_x(x, y)$ и $f'_y(x, y)$ в точке (x, y) и требуется вычислить значение этой функции в точке $(x + \Delta x, y + \Delta y)$.

По определению $f(x + \Delta x, y + \Delta y) - f(x, y) = \Delta z$.

Заменяя Δz полным дифференциалом dz , получаем:

$$f(x + \Delta x, y + \Delta y) - f(x, y) \approx f'_x(x, y) \Delta x + f'_y(x, y) \Delta y,$$

откуда:

$$f(x + \Delta x, y + \Delta y) \approx f(x, y) + f'_x(x, y) \Delta x + f'_y(x, y) \Delta y.$$

Для достаточно малых Δx и Δy допускаемая нами погрешность может быть сделана сколь угодно малой.

Рассмотрим пример. Вычислить приближенно $\sqrt[3]{0,98} \sqrt[4]{1,05^3}$. Рассмотрим функцию $f(x, y) = \sqrt[3]{x} \sqrt[4]{y^3} = x^{\frac{1}{3}} y^{\frac{3}{4}}$; ее частные производные $f'_x(x, y) = \frac{1}{3} x^{-\frac{2}{3}} y^{\frac{3}{4}}$, $f'_y(x, y) = \frac{1}{4} x^{\frac{1}{3}} y^{-\frac{1}{4}}$ в точке $x = 1, y = 1$ непрерывны, т. е. функция дифференцируема в этой точке.

Полагая $\Delta x = -0,02$, $\Delta y = 0,05$, имеем:

$$\begin{aligned} f(x + \Delta x, y + \Delta y) &= \sqrt[3]{0,98} \sqrt[4]{1,05^3} \approx \sqrt[3]{1} \sqrt[4]{1^3} + \frac{1}{3} \cdot 1^{-\frac{2}{3}} \cdot 1^{\frac{3}{4}} (-0,02) + \\ &+ \frac{1}{4} \cdot 1^{\frac{1}{3}} \cdot 1^{-\frac{1}{4}} (0,05) = 1 - \frac{1}{3} 0,02 + \frac{1}{4} 0,05 \approx 1,006. \end{aligned}$$

Данное выше определение полного дифференциала легко обобщается на случай дифференцируемой функции любого числа переменных. *Полным дифференциалом* такой функции в данной точке называется главная, линейная относительно приращений всех аргументов часть полного приращения функции.

Так, например, если $u = x^2 + y^2 + xyz^2$, то в любой точке (x, y, z) :

$$du = (2x + yz^2) \Delta x + (2y + xz^2) \Delta y + 2xyz \Delta z.$$

§ 7. Касательная плоскость и нормаль к поверхности, заданной уравнением $z = f(x, y)$. Геометрический смысл полного дифференциала для функции двух переменных

Существует несколько эквивалентных между собой определений касательной плоскости к поверхности. Предлагаемое ниже определение является естественным обобщением определения касательной (прямой) к кривой.

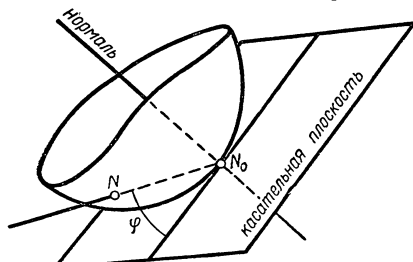


Рис. 303.

Пусть N_0 — точка данной поверхности. Рассмотрим на поверхности другую, переменную, точку N и проведем секущую прямую N_0N .

Плоскость, проходящая через точку N_0 , называется *касательной плоскостью* к поверхности в точке N_0 , если угол между секущей N_0N и этой плоскостью стремится к нулю, когда расстояние N_0N стремится к нулю, каким бы образом точка N на поверхности ни стремилась к точке N_0 (см. рис. 303).

Нормалью к поверхности в точке N_0 называется прямая, проходящая через точку N_0 перпендикулярно касательной плоскости к поверхности в этой точке.

Из определения следует, что у поверхности в данной точке либо есть и тогда только одна касательная плоскость, либо ее нет совсем.

Так, например, поверхность, заданная уравнением $z = \sqrt{x^2 + y^2}$ (коническая поверхность), в точке $O(0, 0, 0)$ касательной плоскости не имеет.

Покажем, что у поверхности, заданной уравнением $z = f(x, y)$, где $f(x, y)$ — функция, дифференцируемая в точке $M_0(x_0, y_0)$, касательная плоскость в точке $N_0(x_0, y_0, f(x_0, y_0))$ существует и имеет уравнение

$$z - f(x_0, y_0) = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0). \quad (1)$$

Пусть $N(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$ — текущая точка поверхности. Обозначим через φ угол между секущей N_0N и плоскостью (1). Покажем, что при стремлении точки N к точке N_0 угол φ или, что равносильно, $\sin \varphi$ стремится к нулю. Этим и будет доказано, что плоскость (1) является касательной к данной поверхности в точке N_0 .

Опустим из точки N перпендикуляр NK на плоскость (1) и перпендикуляр NM на плоскость XOY . $N_1(x_0 + \Delta x, y_0 + \Delta y, z)$ — точка пересечения перпендикуляра NM с плоскостью (1) (рис. 304). Тогда $\angle \varphi = \angle KN_0N$, $|\sin \varphi| = \frac{NK}{N_0N} < \frac{NN_1}{M_0M}$. При этом

$$M_0M = \sqrt{(\Delta x)^2 + (\Delta y)^2} = \rho,$$

$$NN_1 = |f(x_0 + \Delta x, y_0 + \Delta y) - z_{N_1}| = |f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) - f'_x(x_0, y_0)\Delta x - f'_y(x_0, y_0)\Delta y| = |\Delta z - f'_x(x_0, y_0)\Delta x - f'_y(x_0, y_0)\Delta y|.$$

Если точка N стремится к точке N_0 , Δx и Δy , а значит, ρ стремятся к нулю.

Так как функция $f(x, y)$ дифференцируема в точке (x_0, y_0) , величина NN_1 будет бесконечно малой более высокого порядка, чем ρ , т. е. отношение $\frac{NN_1}{M_0M}$ при $N \rightarrow N_0$ будет стремиться к нулю. Отсюда следует, что $\sin \varphi$ и сам угол φ стремятся к нулю, когда $N \rightarrow N_0$.

Таким образом, мы доказали, что если функция $z = f(x, y)$ в точке (x_0, y_0) дифференцируема, то изображающая ее поверхность в точке $N_0(x_0, y_0, f(x_0, y_0))$ имеет невертикальную касательную плоскость.

Можно доказать и обратное утверждение: если в точке

$$N_0(x_0, y_0, f(x_0, y_0))$$

поверхность, изображающая непрерывную функцию $z = f(x, y)$, имеет невертикальную касательную плоскость, то функция $f(x, y)$ в точке (x, y) дифференцируема.

По виду уравнения (1) касательной плоскости к поверхности, заданной уравнением $z = f(x, y)$, в точке N_0 , легко написать уравнения нормали:

$$\frac{x - x_0}{f'_x(x_0, y_0)} = \frac{y - y_0}{f'_y(x_0, y_0)} = \frac{z - z_0}{-1}.$$

Составим, например, уравнения касательной плоскости и нормали к поверхности, заданной уравнением $z = \sqrt{25 - x^2 - y^2}$ в точке $N_0(2, 3, 2\sqrt{3})$.

Так как частные производные

$$\frac{\partial z}{\partial x} = -\frac{x}{\sqrt{25 - x^2 - y^2}} \quad \text{и} \quad \frac{\partial z}{\partial y} = -\frac{y}{\sqrt{25 - x^2 - y^2}}$$

непрерывны в точке $M_0(2, 3)$ и ее окрестности, то функция z дифференцируема в точке M_0 , т. е. данная поверхность имеет в точке N_0 касательную плоскость и нормаль.

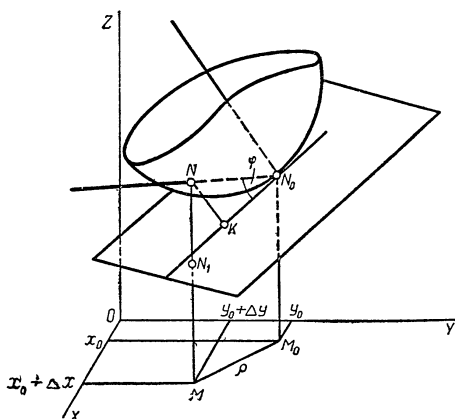


Рис. 304.

Уравнение касательной плоскости:

$$z - 2\sqrt{3} = -\frac{\sqrt{3}}{3}(x - 2) - \frac{\sqrt{3}}{2}(y - 3),$$

уравнения нормали:

$$\frac{x-2}{-\frac{\sqrt{3}}{3}} = \frac{y-3}{-\frac{\sqrt{3}}{2}} = \frac{z-2\sqrt{3}}{-1}, \text{ т. е. } \frac{x-2}{2\sqrt{3}} = \frac{y-3}{3\sqrt{3}} = \frac{z-2\sqrt{3}}{6}.$$

Вясним теперь геометрический смысл полного дифференциала функции двух переменных.

Пусть функция $z = f(x, y)$ дифференцируема в точке (x_0, y_0) . Это значит, что поверхность, заданная уравнением $z = f(x, y)$, имеет в точке $N_0(x_0, y_0, f(x_0, y_0))$ касательную плоскость. Ее уравнение можно записать в виде:

$$z - f(x_0, y_0) = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0),$$

или, полагая $x - x_0 = \Delta x$, $y - y_0 = \Delta y$, в виде:

$$z - f(x_0, y_0) = f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y.$$

В этом равенстве слева стоит разность аппликат точек касательной плоскости, соответствующих точкам (x_0, y_0) и $(x_0 + \Delta x, y_0 + \Delta y)$, а справа — полный дифференциал функции $z = f(x, y)$ в точке (x_0, y_0) .

Таким образом, полный дифференциал функции $z = f(x, y)$ в точке (x_0, y_0)

геометрически обозначает приращение аппликаты касательной плоскости к поверхности, изображающей функцию, в точке $(x_0, y_0, f(x_0, y_0))$ при переходе из точки (x_0, y_0) в точку $(x_0 + \Delta x, y_0 + \Delta y)$ ¹⁾.

Для функции $z = f(x, y)$, изображенной на рис 305, дифференциал dz в точке M_0 отрицателен.

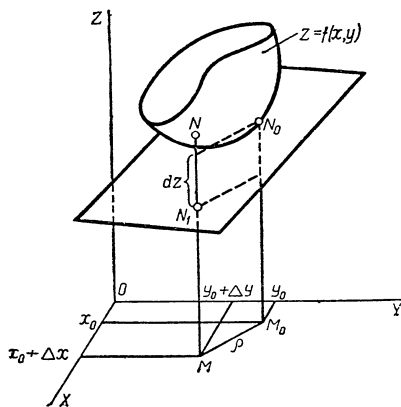


Рис. 305.

§ 8. Производные сложных функций

Пусть $z = f(x, y)$ — функция двух переменных x и y , каждая из которых, в свою очередь, является функцией независимой переменной t : $x = x(t)$, $y = y(t)$. Тогда $z = f(x(t), y(t))$ — сложная функция независимой переменной t ; переменные x и y будут для нее промежуточными переменными.

Например, $z = \frac{x}{y} \arcsin \frac{x-y}{y}$, где $x = 1 - t$, $y = 1 + t^2$, — сложная функция независимой переменной t .

Теорема. Если функции $x(t)$ и $y(t)$ дифференцируемы в точке t , а функция $z = f(x, y)$ дифференцируема в точке (x, y) , где $x = x(t)$, $y = y(t)$, то сложная функция $z = f(x(t), y(t))$ также дифференцируема в точке t , причем

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}. \quad (1)$$

в этой точке.

¹⁾ Напомним, что для функции одной переменной $y = f(x)$ дифференциал в точке x_0 представляет собой приращение ординаты касательной к кривой, изображающей функцию, в точке $(x_0, f(x_0))$ при переходе из точки x_0 в точку $x_0 + \Delta x$.

Доказательство. Пусть приращению Δt переменной t в данной точке t соответствуют приращения Δx и Δy переменных x и y :

$$\Delta x = x(t + \Delta t) - x(t), \quad \Delta y = y(t + \Delta t) - y(t).$$

Приращениям Δx и Δy переменных x и y соответствует приращение функции z :

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

Так как эта функция дифференцируема в точке (x, y) , где $x = x(t)$, $y = y(t)$, имеем:

$$\Delta z = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y + \alpha_1 \Delta x + \alpha_2 \Delta y, \quad (2)$$

где α_1 и α_2 — бесконечно малые при $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$.

Разделив на Δt , получим:

$$\frac{\Delta z}{\Delta t} = f'_x(x, y) \frac{\Delta x}{\Delta t} + f'_y(x, y) \frac{\Delta y}{\Delta t} + \alpha_1 \frac{\Delta x}{\Delta t} + \alpha_2 \frac{\Delta y}{\Delta t}. \quad (3)$$

По условию $\lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} = \frac{dx}{dt}$, $\lim_{\Delta t \rightarrow 0} \frac{\Delta y}{\Delta t} = \frac{dy}{dt}$ в точке t .

Кроме того, так как функции $x(t)$ и $y(t)$ дифференцируемы в точке t , то они непрерывны в этой точке, а значит, при $\Delta t \rightarrow 0$ мы имеем $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$ и, как следствие, $\alpha_1 \rightarrow 0$, $\alpha_2 \rightarrow 0$.

Таким образом, при $\Delta t \rightarrow 0$ существует предел выражения, стоящего в правой части равенства (3), а следовательно, существует $\frac{dz}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta z}{\Delta t}$, причем

$$\frac{dz}{dt} = f'_x(x, y) \frac{dx}{dt} + f'_y(x, y) \frac{dy}{dt},$$

или, короче,

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}.$$

Например, если $z = x \sin \frac{x}{y}$, где $x = 1 + 3t$, $y = \sqrt{1 + t^2}$, то для любого значения t имеем:

$$\frac{dz}{dt} = \left(\sin \frac{x}{y} + \frac{x}{y} \cos \frac{x}{y} \right) \cdot 3 - \frac{x^2}{y^3} \cos \frac{x}{y} \cdot \frac{t}{\sqrt{1 + t^2}}.$$

В частности, если $z = f(x, y)$, где $y = \varphi(x)$, то $z = f(x, \varphi(x))$ — сложная функция x . При выполнении соответствующих условий на основании формулы (2) получим:

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \frac{dy}{dx}.$$

Производная $\frac{\partial z}{\partial x}$, стоящая справа, есть частная производная функции $z = f(x, y)$ по x . Она вычисляется при закреплённом значении y . Производная $\frac{dz}{dx}$ в левой части равенства есть *полная производная* z как сложной функции x , зависящей от x явным образом и через посредство переменной y .

Например, если $z = x^2 y^5 + e^{xy}$, где $y = \frac{1}{1 + x^2}$, то для любого значения x имеем:

$$\frac{dz}{dx} = 2xy^5 + ye^{xy} + (5x^2 y^4 + xe^{xy}) \frac{-2x}{(1 + x^2)^2}.$$

Аналогичным рассмотренному выше образом решается вопрос о производной $\frac{du}{dt}$, когда u задается как дифференцируемая функция любого числа переменных, каждая из которых, в свою очередь, есть дифференцируемая функция t .

Так, в случае $u = f(x, y, z)$, где $x = x(t)$, $y = y(t)$, $z = z(t)$, при выполнении соответствующих условий находим:

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt} + \frac{\partial u}{\partial z} \frac{dz}{dt}.$$

Рассмотрим теперь более общий случай. Пусть $z = f(x, y)$ — функция двух переменных, аргументы которой x и y , в свою очередь, зависят от двух или большего числа переменных. Например, пусть

$$x = x(u, v), \quad y = y(u, v).$$

Тогда $z = f(x(u, v), y(u, v))$ будет *сложной функцией* независимых переменных u и v .

Если функции $x(u, v)$ и $y(u, v)$ дифференцируемы в точке (u, v) , а функция $z = f(x, y)$ дифференцируема в точке (x, y) , где $x = x(u, v)$, $y = y(u, v)$, то сложная функция $z = f(x(u, v), y(u, v))$ дифференцируема в точке (u, v) ¹⁾, причем ее частные производные в этой точке находятся по формулам:

$$\begin{aligned} \frac{\partial z}{\partial u} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u}, \\ \frac{\partial z}{\partial v} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v}. \end{aligned}$$

Эти формулы получаем как следствие формулы (2), так как при нахождении $\frac{\partial z}{\partial u} \left(\frac{\partial z}{\partial v} \right)$ z рассматривается как функция одной переменной $u(v)$ при постоянном значении второй переменной.

К аналогичным выводам мы приходим и в случаях, когда число промежуточных аргументов или число независимых аргументов больше двух.

§ 9. Инвариантная форма записи дифференциала

По определению, для любой дифференцируемой в точке (x, y) функции $z = f(x, y)$

двух независимых переменных

$$dz = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y. \quad (1)$$

Полагая, в частности, $z = x$ (т. е. $f(x, y) \equiv x$), получаем $dz = dx = 1 \cdot \Delta x + 0 \cdot \Delta y$, т. е. $dx = \Delta x$.

Аналогично, полагая $z = y$, получаем, что $dy = \Delta y$.

Таким образом, дифференциалы независимых переменных совпадают с приращениями этих переменных, и мы можем записать дифференциал функции $f(x, y)$ в виде:

$$dz = f'_x(x, y) dx + f'_y(x, y) dy$$

или, что то же, в виде:

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy. \quad (2)$$

Эта форма записи дифференциала называется *инвариантной*, что объясняется приведенными ниже соображениями.

Пусть $z = f(x, y)$, где $x = x(u, v)$, $y = y(u, v)$, — сложная функция независимых переменных u и v .

Предположим, что функции $x = x(u, v)$ и $y = y(u, v)$ дифференцируемы в точке (u, v) , а функция $f(x, y)$ дифференцируема в точке (x, y) , где $x = x(u, v)$, $y = y(u, v)$.

¹⁾ Доказательство дифференцируемости z мы опускаем, предоставляя его читателю.

Тогда сложная функция

$$z = f(x(u, v), y(u, v))$$

будет дифференцируема в точке (u, v) .

При этом, согласно (2), $dz = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv$.

Применяя правила для вычисления частных производных сложной функции, получаем:

$$\begin{aligned} dz &= \left(\frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u} \right) du + \left(\frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v} \right) dv = \\ &= \frac{\partial z}{\partial x} \left(\frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv \right) + \frac{\partial z}{\partial y} \left(\frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv \right). \end{aligned}$$

Так как в скобках стоят полные дифференциалы функций $x(u, v)$, $y(u, v)$, то окончательно имеем:

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy.$$

Таким образом, и в том случае, когда x и y — независимые переменные, и в том случае, когда x и y — зависимые переменные, дифференциал функции $z = f(x, y)$ можно записать в форме:

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy.$$

В связи с этим данная форма записи полного дифференциала и называется инвариантной.

Предложенная ранее форма записи полного дифференциала

$$dz = \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y$$

не будет инвариантной, она может использоваться только, если x и y — независимые переменные, так как в противном случае

$$\Delta x \neq dx, \quad \Delta y \neq dy.$$

Пользуясь определением полного дифференциала и инвариантной формой записи, можно получить ряд формул, аналогичных соответствующим формулам для функций одной переменной.

Пусть $z = f(x, y)$ и $z = \varphi(x, y)$ — дифференцируемые функции двух (независимых или зависимых) переменных x и y в некоторой точке (x, y) .

Тогда в этой точке $d(Cf) = Cdf$ (C — число), $d(f \pm \varphi) = df \pm d\varphi$, $d(f \cdot \varphi) = df \cdot \varphi + f \cdot d\varphi$, $d \frac{f}{\varphi} = \frac{df \cdot \varphi - f \cdot d\varphi}{\varphi^2}$.

Для функции любого числа переменных $u = f(x, y, z, \dots, w)$, отождествляя приращения независимых переменных с их дифференциалами, также можно дать так называемую *инвариантную форму* записи полного дифференциала:

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz + \dots + \frac{\partial u}{\partial w} dw.$$

Она будет справедливой и в том случае, когда $x, y, z, \dots, w$ будут не независимыми, а промежуточными аргументами функции u .

§ 10. Производная по направлению

Рассмотрим функцию

$$z = f(x, y) \quad (1)$$

в некоторой области, содержащей точку (x, y) , и единичный вектор $\mathbf{n} \{\cos \alpha, \cos \beta\}$ любого направления. Для характеристики скорости изменения функции в точке (x, y) в направлении $\mathbf{n}$ введем понятие производной по направлению.

Перейдем из точки $M(x, y)$ в точку $M_1(x + \Delta x, y + \Delta y)$ в направлении вектора $\mathbf{n}$ (рис. 306). Функция $z = f(x, y)$ получит при этом приращение

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

Разделив приращение функции на длину отрезка смещения MM_1 , получим отношение $\frac{\Delta z}{MM_1}$.

Предел этого отношения при стремлении точки M к точке M_1 , если он существует и конечен, и называется *производной* функции $z = f(x, y)$ в точке M по направлению вектора $\mathbf{n}$.

Обозначается производная по направлению $\mathbf{n}$ следующим образом: $\frac{\partial f(x, y)}{\partial n}$ или $\frac{\partial z}{\partial n}$. Таким образом, по определению

$$\frac{\partial z}{\partial n} = \lim_{MM_1 \rightarrow 0} \frac{\Delta z}{MM_1}.$$

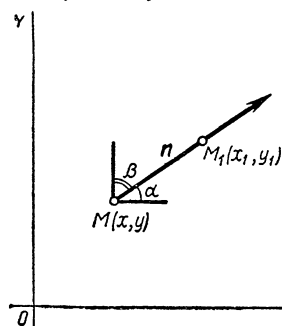


Рис. 306.

Так как отношение $\frac{\Delta z}{MM_1}$ дает среднюю скорость изменения функции z на участке MM_1 , то его предел при $MM_1 \rightarrow 0$ естественно принять за истинную скорость изменения функции z в точке M в направлении вектора $\mathbf{n}$.

Обозначая через $\mathbf{i}$ и $\mathbf{j}$ единичные векторы в направлении осей OX и OY , видим, что если в данной точке существуют частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, то, полагая $\mathbf{n} = \mathbf{i}$, получим $\frac{\partial z}{\partial n} = \frac{\partial z}{\partial x}$, а полагая $\mathbf{n} = \mathbf{j}$, получим $\frac{\partial z}{\partial n} = \frac{\partial z}{\partial y}$.

Предположим теперь, что функция $z = f(x, y)$ дифференцируема в точке $M(x, y)$. Обозначая MM_1 через ρ , имеем:

$$\Delta x = \rho \cos \alpha, \quad \Delta y = \rho \cos \beta,$$

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y) = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y + \tilde{\alpha} \rho,$$

т. е.

$$\Delta z = f'_x(x, y) \rho \cos \alpha + f'_y(x, y) \rho \cos \beta + \tilde{\alpha} \rho, \quad (2)$$

где $\tilde{\alpha}$ — бесконечно малая при $\rho \rightarrow 0$. Деля на ρ и переходя к пределу при $\rho \rightarrow 0$, получаем:

$$\frac{\partial z}{\partial n} = f'_x(x, y) \cos \alpha + f'_y(x, y) \cos \beta$$

или, короче,

$$\frac{\partial z}{\partial n} = \frac{\partial z}{\partial x} \cos \alpha + \frac{\partial z}{\partial y} \cos \beta. \quad (3)$$

Таким образом, для дифференцируемой функции знание частных производных (т. е. производных по двум взаимно перпендикулярным направлениям) позволяет непосредственно написать выражение производной по любому направлению.

Пример. Вычислить производную от функции $z = x^2 + y^2x$ в точке $A(1, 2)$ по направлению вектора $\overrightarrow{AB}$, где B — точка $(3, 0)$.

Решение. Находим единичный вектор $\mathbf{n}$ данного направления:

$$\overrightarrow{AB} = 2\mathbf{i} - 2\mathbf{j}, \quad |\overrightarrow{AB}| = \sqrt{8} = 2\sqrt{2},$$

$$\mathbf{n} = \frac{\overrightarrow{AB}}{|\overrightarrow{AB}|} = \frac{1}{\sqrt{2}} \mathbf{i} - \frac{1}{\sqrt{2}} \mathbf{j}, \quad \text{т. е. } \cos \alpha = \frac{1}{\sqrt{2}}, \quad \cos \beta = -\frac{1}{\sqrt{2}},$$

затем вычисляем частные производные $\frac{\partial z}{\partial x} = 2x + y^2$ и $\frac{\partial z}{\partial y} = 2xy$.

Так как в точке A $\frac{\partial z}{\partial x} = 6$, $\frac{\partial z}{\partial y} = 4$, то по формуле (3) получаем

$$\frac{\partial z}{\partial n} = 6 \cdot \frac{1}{\sqrt{2}} - 4 \cdot \frac{1}{\sqrt{2}} = \sqrt{2}.$$

Из формулы (3) следует, что при замене направления $\mathbf{n}$ на противоположное производная по направлению (для дифференцируемой функции) меняет знак.

Вектор с координатами $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$ в точке (x, y) называется *градиентом* функции $z = f(x, y)$ в этой точке и обозначается $\text{grad } f(x, y)$ или $\text{grad } z$.

По определению

$$\text{grad } z = \frac{\partial z}{\partial x} \mathbf{i} + \frac{\partial z}{\partial y} \mathbf{j}.$$

Записывая формулу (3) с помощью скалярного произведения, имеем для любого направления $\mathbf{n}$

$$\frac{\partial z}{\partial n} = (\text{grad } z, \mathbf{n}) = n_p \text{ grad } z = |\text{grad } z| \cdot \cos \omega,$$

где ω — угол между направлением $\mathbf{n}$ и направлением градиента в данной точке.

Из полученного соотношения следует, что производная функции по направлению будет наибольшей, если это направление совпадает с направлением градиента ($\omega = 0$). При этом $\frac{\partial z}{\partial n} = |\text{grad } z|$.

Таким образом, градиент дифференцируемой функции $z = f(x, y)$ в каждой точке по величине и по направлению дает наибольшую скорость изменения функции в этой точке.

Аналогичным образом определяется *производная по направлению* для функции трех переменных $u = f(x, y, z)$, выводится формула, связывающая производную по направлению $\mathbf{n}\{\cos \alpha, \cos \beta, \cos \gamma\}$ с частными производными $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial u}{\partial z}$ дифференцируемой функции:

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma,$$

вводится понятие *градиента* функции и исследуются его свойства.

§ 11. Неявные функции

В гл. III, § 2 мы ввели понятие неявной функции, заданной уравнением

$$F(x, y) = 0. \quad (1)$$

Повторим вкратце основные определения.

Рассмотрим множество тех значений переменной x , для которых существуют значения переменной y , совместно с x обращающие уравнение (1) в тождество.

Если каждому значению x из этого множества поставить в соответствие те значения y , для которых $F(x, y) = 0$, мы получим однозначную или многозначную функцию $y = f(x)$. Эта функция называется *неявно заданной* уравнением $F(x, y) = 0$, или, короче, *неявной функцией*.

Рассмотренное выше множество значений x является областью определения этой функции. Для любого значения x из области определения неявной функции имеем $F(x, f(x)) = 0^1$.

¹⁾ Т. е. для любого значения x из рассматриваемого множества $y = f(x)$ будет по определению решением уравнения $F(x, y) = 0$.

Так, например, уравнение $x^2 + y^2 - 1 = 0$ определяет неявную функцию в области $-1 \leq x \leq 1$: каждому значению x из этой области соответствуют два значения y ($y = \pm \sqrt{1 - x^2}$ и $y = -\sqrt{1 - x^2}$), таких, что $x^2 + y^2 - 1 = 0$.

Уравнение $\sin(xy) = 0$ неявно задает y как бесконечнозначную функцию x на всей числовой оси. При этом $x = 0$ соответствуют все значения y , а любому $x \neq 0$ — значения

$$y = 0, \pm \frac{\pi}{x}, \pm \frac{\pi}{2x}, \dots$$

Однако не всякое уравнение, связывающее переменные x и y , определяет неявную функцию.

Уравнения $x^2 + y^2 + 4 = 0$, $\sin(x + y) - 5 = 0$, $e^{xy} = 0$ и т. д. не удовлетворяются ни одной парой действительных чисел x и y , т. е. не определяют никаких неявных функций.

Возникает задача о нахождении условий, накладываемых на функцию двух переменных $F(x, y)$, при которых уравнение $F(x, y) = 0$ определяет неявную функцию $y = f(x)$. *Достаточные условия существования неявной функции*, удобные для применения, содержатся в следующей теореме, которую мы приводим без доказательства.

Теорема 1 (теорема существования неявной функции).

Пусть:

1) функция $F(x, y)$ определена и непрерывна вместе со своими частными производными $F'_x(x, y)$ и $F'_y(x, y)$ в некоторой окрестности точки (x_0, y_0) ,

2) $F(x, y)$ в точке (x_0, y_0) обращается в нуль: $F(x_0, y_0) = 0$,

3) $F'_y(x, y)$ в точке (x_0, y_0) отлична от нуля: $F'_y(x_0, y_0) \neq 0$.

Тогда:

1) в некотором прямоугольнике $D \left\{ \begin{array}{l} x_0 - \delta < x < x_0 + \delta, \\ y_0 - \Delta < y < y_0 + \Delta \end{array} \right\}$ уравнение $F(x, y) = 0$ определяет y как однозначную функцию от x : $y = f(x)$,

2) при $x = x_0$ эта функция принимает значение y_0 : $f(x_0) = y_0$,

3) на интервале $x_0 - \delta < x < x_0 + \delta$ функция $y = f(x)$ непрерывна и имеет непрерывную производную.

Геометрически это значит, что кривая, заданная «неявным» уравнением $F(x, y) = 0$, в прямоугольнике D представляет собой проходящий через точку (x_0, y_0) график однозначной, непрерывной и непрерывно-дифференцируемой функции $y = f(x)$.

Функция $y = f(x)$, неявно заданная уравнением $F(x, y) = 0$, согласно теореме существования дифференцируема в окрестности $(x_0 - \delta, x_0 + \delta)$ точки x_0 . Найдем ее производную.

Так как в этой окрестности $F(x, f(x)) \equiv 0$, то для любой ее точки x $\frac{d}{dx} F(x, f(x)) = 0$ или, что то же самое, $\frac{dF(x, y)}{dx} = 0$, где $y = f(x)$.

Вычисляя полную производную, получаем:

$$F'_x(x, y) + F'_y(x, y) y' = 0,$$

откуда

$$y' = - \frac{F'_x(x, y)}{F'_y(x, y)} \quad (2)$$

Таким образом, мы можем найти производную неявной функции, не имея непосредственного явного задания самой функции.

Пример. Дано уравнение $x^4 + y^5 + e^{xy} = 0$. Очевидно, оно удовлетворяется при $x_0 = 0$, $y_0 = -1$.

¹⁾ Функция $F'_y(x, y)$ в окрестности точки (x_0, y_0) непрерывна, $F'_y(x_0, y_0) \neq 0$, следовательно, $F'_y(x, y) \neq 0$ и в некоторой окрестности точки (x_0, y_0) .

Функции $F(x, y) = x^4 + y^5 + e^{xy}$, $F'_x(x, y) = 4x^3 + ye^{xy}$, $F'_y(x, y) = 5y^4 + xe^{xy}$ непрерывны в окрестности точки $(0, -1)$, причем $F'_y(0, -1) = -5 \neq 0$.

По теореме существования можно утверждать, что в некотором прямоугольнике $\{-\delta < x < \delta, -1 - \Delta < y < -1 + \Delta\}$ существует однозначная функция $y = f(x)$, определяемая данным уравнением, такая, что $f(0) = -1$.

Эта функция в окрестности $-\delta < x < \delta$ непрерывна и имеет непрерывную производную

$$y' = -\frac{4x^3 + ye^{xy}}{5y^4 + xe^{xy}}.$$

В частности, в точке $(0, -1)$

$$y' = -\frac{-1}{5} = \frac{1}{5}.$$

Пусть теперь дано уравнение

$$F(x, y, z) = 0, \quad (3)$$

связывающее значения трех переменных x, y и z . Рассмотрим множество тех пар чисел (x, y) , для которых существуют значения z , обращающие совместно с x, y уравнение (3) в тождество.

Подставив каждой паре чисел (x, y) из этого множества в соответствие те значения z , для которых $F(x, y, z) = 0$, мы получим однозначную или многозначную функцию двух переменных:

$$z = f(x, y).$$

Областью определения этой функции будет множество всех пар чисел (x, y) (всех точек (x, y) плоскости XOY), для которых

$$F(x, y, f(x, y)) = 0.$$

Функцию $z = f(x, y)$ мы будем называть *неявно заданной* уравнением (3), или просто *неявной функцией*¹⁾.

Примеры неявно заданных функций:

1. Уравнение $x + y + z + 1 = 0$ определяет z как неявную функцию x и y во всей плоскости XOY . Каждой паре чисел (x, y) соответствует одно значение z : $z = -1 - x - y$, такое, что

$$x + y + z + 1 = 0.$$

Явное выражение данной функции: $z = -1 - x - y$.

2. Уравнение $x^2 + y^2 + z^2 - 1 = 0$ задает z как неявную функцию x и y в круге $x^2 + y^2 \leq 1$. Эта функция двузначна. Ее явное выражение $z = \pm \sqrt{1 - x^2 - y^2}$.

3. Уравнение $yz = x \ln yz$ задает z как неявную функцию x, y . Можно показать, что эта функция не может быть задана в явной форме с помощью конечного числа элементарных функций.

Уравнение $F(x, y, z) = 0$ не всегда определяет z как неявную функцию x и y . Так, например, уравнение $\sin(x + y + 3z) = 4$ не удовлетворяется ни одной тройкой чисел x, y, z .

Можно, однако, показать, что если в некоторой окрестности точки (x_0, y_0, z_0) функция $F(x, y, z)$ и ее частные производные $F'_x(x, y, z)$, $F'_y(x, y, z)$, $F'_z(x, y, z)$ непрерывны, причем $F(x_0, y_0, z_0) = 0$, $F'_z(x_0, y_0, z_0) \neq 0$, то существует параллелепипед

$$D \{ x_0 - \delta < x < x_0 + \delta, y_0 - \delta < y < y_0 + \delta, z_0 - \Delta < z < z_0 + \Delta \},$$

в котором уравнение (3) определяет z как однозначную непрерывную и

¹⁾ Для любой пары чисел (x, y) из рассматриваемого множества по определению $z = f(x, y)$ есть решение уравнения $F(x, y, z) = 0$.

имеющую непрерывные частные производные функцию от x и y : $z = f(x, y)$, удовлетворяющую условию $f(x_0, y_0) = z_0$.

Частные производные такой функции вычисляются по формулам:

$$\frac{\partial z}{\partial x} = -\frac{F'_x(x, y, z)}{F'_z(x, y, z)}, \quad \frac{\partial z}{\partial y} = -\frac{F'_y(x, y, z)}{F'_z(x, y, z)}.$$

Эти функции можно получить как следствие формулы (2), так как при вычислении частных производных $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ переменные y и x считаются соответственно постоянными.

Геометрически высказанное утверждение означает, что поверхность, заданная уравнением $F(x, y, z) = 0$, в параллелепипеде D может быть задана уравнением $z = f(x, y)$, где $f(x, y)$ однозначная, непрерывная и имеющая непрерывные частные производные функция x и y . Точка (x_0, y_0, z_0) , где $z_0 = f(x_0, y_0)$, лежит на этой поверхности.

Рассматривая уравнение $F(x, y, z, \dots, w) = 0$, связывающее значения n переменных, мы можем аналогично тому, как делалось выше, ввести понятие неявной функции от $n - 1$ переменных, сформулировать условия ее существования, дать формулы для вычисления частных производных.

Общая теория неявных функций изучает и более широкую постановку вопроса, когда дается не одно, а несколько уравнений:

$$\begin{aligned} F_1(x, y, z, \dots, w) &= 0, \\ F_2(x, y, z, \dots, w) &= 0, \\ &\dots\dots\dots \\ F_m(x, y, z, \dots, w) &= 0, \end{aligned}$$

связывающих значения n переменных $x, y, z, \dots, w, n > m$.

При определенных условиях, наложенных на функции $F_1, F_2, \dots, F_m$, эту систему уравнений можно рассматривать как неявное задание m функций от $n - m$ независимых переменных.

Остановимся подробнее на простейшем случае — системе из двух уравнений с тремя переменными:

$$\begin{cases} F_1(x, y, z) = 0, \\ F_2(x, y, z) = 0. \end{cases} \quad (4)$$

Рассмотрим множество всех значений x , для которых существуют пары чисел y и z , обращающих совместно с x уравнения (4) в тождества.

Поставив каждому значению x из этого множества в соответствие значения y и z , для которых $F_1(x, y, z) = 0$ и $F_2(x, y, z) = 0$, мы получим две функции $y = f_1(x)$ и $z = f_2(x)$, которые *неявно задаются* системой (4).

Имеет место следующая теорема:

Теорема 2. Если:

1) в некоторой окрестности точки (x_0, y_0, z_0) функции $F_1(x, y, z)$ и $F_2(x, y, z)$ непрерывны и имеют непрерывные частные производные по x, y и z ,

2) в точке (x_0, y_0, z_0) эти функции обращаются в нуль:

$$F_1(x_0, y_0, z_0) = 0 \text{ и } F_2(x_0, y_0, z_0) = 0,$$

3) в точке (x_0, y_0, z_0) определитель

$$\begin{vmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{vmatrix}$$

(обозначим его $\frac{\partial(F_1, F_2)}{\partial(y, z)}$) отличен от нуля,

то:

1) в некотором параллелепипеде

$$D \left\{ \begin{aligned} x_0 - \delta &< x < x_0 + \delta, \\ y_0 - \Delta &< y < y_0 + \Delta, \\ z_0 - \Delta &< z < z_0 + \Delta \end{aligned} \right\}$$

система (4) определяет y и z как однозначные функции от x :

$$y = f_1(x), \quad z = f_2(x),$$

2) при $x = x_0$ эти функции принимают соответственно значения y_0 и z_0 :

$$y_0 = f_1(x_0), \quad z_0 = f_2(x_0),$$

3) на интервале $x_0 - \delta < x < x_0 + \delta$ эти функции $f_1(x)$ и $f_2(x)$ непрерывны и имеют непрерывные производные.

Выведем формулы для вычисления этих производных.

Если $y = f_1(x)$, $z = f_2(x)$, то в интервале $(x_0 - \delta, x_0 + \delta)$ равенства $F_1(x, y, z) = 0$, $F_2(x, y, z) = 0$ выполняются тождественно.

Дифференцируя их по x , получим:

$$\frac{\partial F_1}{\partial x} + \frac{\partial F_1}{\partial y} \frac{\partial y}{\partial x} + \frac{\partial F_1}{\partial z} \frac{\partial z}{\partial x} = 0,$$

$$\frac{\partial F_2}{\partial x} + \frac{\partial F_2}{\partial y} \frac{\partial y}{\partial x} + \frac{\partial F_2}{\partial z} \frac{\partial z}{\partial x} = 0.$$

Эта система линейных уравнений относительно неизвестных $\frac{\partial y}{\partial x}$ и $\frac{\partial z}{\partial x}$ с определителем $\frac{\partial(F_1, F_2)}{\partial(y, z)}$, отличным от нуля в окрестности рассматриваемой точки.

Решая систему по формулам Крамера, получим:

$$\frac{\partial y}{\partial x} = -\frac{\frac{\partial(F_1, F_2)}{\partial(x, z)}}{\frac{\partial(F_1, F_2)}{\partial(y, z)}}, \quad \frac{\partial z}{\partial x} = -\frac{\frac{\partial(F_1, F_2)}{\partial(y, x)}}{\frac{\partial(F_1, F_2)}{\partial(y, z)}}.$$

Важно отметить, что все высказанные выше утверждения имеют локальный характер: речь идет все время лишь о некоторой окрестности рассматриваемой точки.

Примечание. Не следует смешивать вопрос о существовании неявных функций с вопросом о явном задании этих функций. Функция может существовать, но не выражаться в явной форме через элементарные функции. Приведенные выше теоремы дают достаточные условия существования неявных функций.

Рассмотрим теперь некоторые приложения теории неявных функций.

1. Касательная и нормаль к плоской кривой, заданной уравнением $F(x, y) = 0$. Пусть плоская кривая задана уравнением $F(x, y) = 0$. Возьмем на ней точку (x_0, y_0) .

Если в окрестности этой точки функция $F(x, y)$ удовлетворяет условиям теоремы 1, то вблизи точки (x_0, y_0) кривая может быть задана уравнением $y = f(x)$, где $f(x)$ — однозначная, непрерывная, имеющая непрерывную производную, функция. Отсюда, в частности, следует, что в точке (x_0, y_0) кривая имеет невертикальную касательную.

Ее уравнение $y - y_0 = \frac{dy}{dx} \Big|_{(x_0, y_0)} (x - x_0)$ или, так как

$$\frac{dy}{dx} \Big|_{(x_0, y_0)} = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)}, \quad \text{то } y - y_0 = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)} (x - x_0).$$

После элементарных преобразований получим уравнение касательной к кривой $F(x, y) = 0$ в точке (x_0, y_0) в виде:

$$F'_x(x_0, y_0)(x - x_0) + F'_y(x_0, y_0)(y - y_0) = 0. \quad (5)$$

Учитывая условия перпендикулярности прямых, запишем уравнение нормали к кривой $F(x, y) = 0$ в точке (x_0, y_0) в виде:

$$F'_y(x_0, y_0)(x - x_0) - F'_x(x_0, y_0)(y - y_0) = 0 \quad \text{или} \quad \frac{x - x_0}{F'_x(x_0, y_0)} = \frac{y - y_0}{F'_y(x_0, y_0)}.$$

Если функция $F(x, y)$ вместо условия $F'_y(x_0, y_0) \neq 0$ удовлетворяет условию $F'_x(x_0, y_0) \neq 0$, то в некоторой окрестности точки (x_0, y_0) кривая $F(x, y) = 0$ может быть задана уравнением $x = \varphi(y)$, где $\varphi(y)$ — одно-

значная непрерывная функция, имеющая непрерывную производную. Это значит, что в точке (x_0, y_0) кривая имеет негоризонтальную касательную. Ее уравнение, как легко видеть, снова запишется в виде (5). Точки, в которых $F'_x(x_0, y_0) = F'_y(x_0, y_0) = 0$, называются *особыми* точками кривой. В особой точке кривая может не иметь определенной касательной.

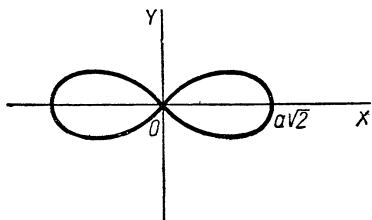


Рис. 307.

Так, например, кривая $(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$ — лемниската Бернулли, в точке $(0, 0)$ касательной не имеет (рис. 307).

Частные производные функции $F(x, y) = (x^2 + y^2)^2 - 2a^2(x^2 - y^2)$:
 $F'_x(x, y) = 2(x^2 + y^2)2x - 4a^2x$ и $F'_y(x, y) = 2(x^2 + y^2)2y + 4a^2y$
 в точке $(0, 0)$ равны нулю, точка $(0, 0)$ — особая точка кривой.

Пример. Найти уравнение касательной и нормали к эллипсу $\frac{x^2}{4} + \frac{y^2}{9} = 1$ в точке $M_0\left(\sqrt{3}, -\frac{3}{2}\right)$.

Здесь $F(x, y) = \frac{x^2}{4} + \frac{y^2}{9} - 1$, $F'_x(x, y) = \frac{x}{2}$, $F'_y(x, y) = \frac{2y}{9}$ — функции, непрерывные всюду.

Точка $\left(\sqrt{3}, -\frac{3}{2}\right)$ — не особая, так как $F'_y\left(\sqrt{3}, -\frac{3}{2}\right) \neq 0$. Кривая имеет в этой точке касательную и нормаль. Их уравнения соответственно запишутся в виде:

$$\frac{\sqrt{3}}{2}(x - \sqrt{3}) - \frac{1}{3}\left(y + \frac{3}{2}\right) = 0 \text{ (уравнение касательной),}$$

$$\frac{x - \sqrt{3}}{\frac{\sqrt{3}}{2}} = \frac{y + \frac{3}{2}}{-\frac{1}{3}} \text{ (уравнение нормали).}$$

2. Касательная плоскость и нормаль к поверхности, заданной уравнением $F(x, y, z) = 0$. Пусть поверхность задана уравнением

$$F(x, y, z) = 0.$$

Возьмем на ней точку (x_0, y_0, z_0) .

Если в окрестности этой точки функция $F(x, y, z)$ и ее частные производные F'_x, F'_y, F'_z непрерывны, причем $F'_z(x_0, y_0, z_0) \neq 0$, то вблизи точки (x_0, y_0, z_0) поверхность можно представить уравнением $z = f(x, y)$, где функция $f(x, y)$ однозначна, непрерывна и имеет непрерывные частные производные. Отсюда, в частности, следует, что в точке (x_0, y_0) функция $z = f(x, y)$ дифференцируема, т. е. данная поверхность имеет неперпендикулярную касательную плоскость.

Уравнение этой плоскости, как мы знаем, имеет вид:

$$z - z_0 = \left. \frac{\partial z}{\partial x} \right|_{(x_0, y_0)} (x - x_0) + \left. \frac{\partial z}{\partial y} \right|_{(x_0, y_0)} (y - y_0).$$

Подставляя в него значения $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$:

$$\left. \frac{\partial z}{\partial x} \right|_{(x_0, y_0)} = -\frac{F'_x(x_0, y_0, z_0)}{F'_z(x_0, y_0, z_0)}, \quad \left. \frac{\partial z}{\partial y} \right|_{(x_0, y_0)} = -\frac{F'_y(x_0, y_0, z_0)}{F'_z(x_0, y_0, z_0)},$$

после элементарных преобразований получим уравнение касательной плоскости к поверхности $F(x, y, z) = 0$ в точке (x_0, y_0, z_0) в виде:

$$F'_x(x_0, y_0, z_0)(x - x_0) + F'_y(x_0, y_0, z_0)(y - y_0) + F'_z(x_0, y_0, z_0)(z - z_0) = 0. \quad (6)$$

Уравнения нормали к той же поверхности в точке (x_0, y_0, z_0) будут иметь вид:

$$\frac{x-x_0}{F'_x(x_0, y_0, z_0)} = \frac{y-y_0}{F'_y(x_0, y_0, z_0)} = \frac{z-z_0}{F'_z(x_0, y_0, z_0)}. \quad (7)$$

Если вместо условия $F'_z(x_0, y_0, z_0) \neq 0$ выполнены условия $F'_x(x_0, y_0, z_0) \neq 0$ или $F'_y(x_0, y_0, z_0) \neq 0$, то, рассуждая аналогичным образом, получим уравнения касательной плоскости и нормали к поверхности в том же виде (6) и (7).

Точка (x_0, y_0, z_0) , в которой $F'_x = F'_y = F'_z = 0$, называется *особой* точкой поверхности. В этой точке поверхность может не иметь касательной плоскости.

Рассмотрим пример. Найти уравнения касательной плоскости и нормали к поверхности $x^2 + y^2 - z^2 = 0$ в точке $N(0, 1, 1)$.

Здесь $F(x, y, z) = x^2 + y^2 - z^2$, $F'_x(x, y, z) = 2x$, $F'_y(x, y, z) = 2y$, $F'_z(x, y, z) = -2z$ — функции, непрерывные всюду, $F'_z(0, 1, 1) = -2 \neq 0$, следовательно, в точке $(0, 1, 1)$ можно провести касательную плоскость и нормаль к поверхности. Подсчитав $F'_x(0, 1, 1) = 0$, $F'_y(0, 1, 1) = 2$, $F'_z(0, 1, 1) = -2$, получаем уравнения искомой касательной плоскости

$$2(y-1) - 2(z-1) = 0, \text{ т. е. } y - z = 0,$$

и нормали

$$\frac{x-0}{0} = \frac{y-1}{2} = \frac{z-1}{-2}.$$

В точке $O(0, 0, 0)$ $F'_x = F'_y = F'_z = 0$, т. е. O — особая точка поверхности. В этой точке касательной плоскости и нормали к данной поверхности не существует (рис. 308).

3. Касательная прямая и нормальная плоскость к пространственной кривой, заданной уравнениями $F_1(x, y, z) = 0$, $F_2(x, y, z) = 0$. При определенных условиях уравнения касательной к пространственной кривой, заданной уравнениями в параметрической форме $x = x(t)$, $y = y(t)$, $z = z(t)$, в точке $M_0(x_0, y_0, z_0)$ имеют вид:

$$\frac{x-x_0}{\frac{dx}{dt}} = \frac{y-y_0}{\frac{dy}{dt}} = \frac{z-z_0}{\frac{dz}{dt}} \quad (\text{см. гл. IV, § 24}).$$

Пусть теперь кривая задана как линия пересечения поверхностей $F_1(x, y, z) = 0$ и $F_2(x, y, z) = 0$. Возьмем на этой кривой точку (x_0, y_0, z_0) . Если в окрестности этой точки выполнены условия теоремы 2, то вблизи точки (x_0, y_0, z_0) кривая может быть задана уравнениями $y = f_1(x)$, $z = f_2(x)$, где $f_1(x)$ и $f_2(x)$ — однозначные, непрерывные, имеющие непрерывные производные функции. Их производные определяются по формулам:

$$\frac{dy}{dx} = -\frac{\frac{\partial(F_1, F_2)}{\partial(x, z)}}{\frac{\partial(F_1, F_2)}{\partial(y, z)}}, \quad \frac{dz}{dx} = -\frac{\frac{\partial(F_1, F_2)}{\partial(y, x)}}{\frac{\partial(F_1, F_2)}{\partial(y, z)}}.$$

Считая x параметром, запишем уравнения касательной к данной кривой в точке (x_0, y_0, z_0) :

$$\frac{x-x_0}{\frac{dx}{dx}} = \frac{y-y_0}{\frac{dy}{dx}} = \frac{z-z_0}{\frac{dz}{dx}}.$$

Подставляя значения производных, вычисленных в точке (x_0, y_0, z_0) , получим уравнения касательной к кривой $\begin{cases} F_1(x, y, z) = 0, \\ F_2(x, y, z) = 0 \end{cases}$ в точке (x_0, y_0, z_0) в виде:

$$\frac{x-x_0}{\frac{\partial(F_1, F_2)}{\partial(y, z)}} = \frac{y-y_0}{\frac{\partial(F_1, F_2)}{\partial(z, x)}} = \frac{z-z_0}{\frac{\partial(F_1, F_2)}{\partial(x, y)}}.$$

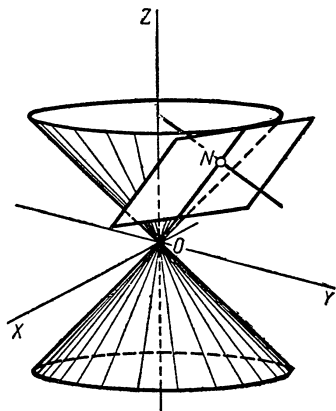


Рис. 308.

Уравнение нормальной плоскости к той же кривой в точке (x_0, y_0, z_0) будет иметь вид:

$$\frac{\partial (F_1, F_2)}{\partial (y, z)} (x - x_0) + \frac{\partial (F_1, F_2)}{\partial (z, x)} (y - y_0) + \frac{\partial (F_1, F_2)}{\partial (x, y)} (z - z_0) = 0.$$

Здесь все коэффициенты вычислены в точке (x_0, y_0, z_0) .

Примечание. Точки, в которых $\frac{\partial (F_1, F_2)}{\partial (y, z)} = \frac{\partial (F_1, F_2)}{\partial (z, x)} = \frac{\partial (F_1, F_2)}{\partial (x, y)} = 0$, называются *особыми* точками кривой. В этих точках кривая может не иметь определенной касательной.

§ 12. Частные производные высших порядков

Если функция $z = f(x, y)$ определена в некоторой области D , то ее частные производные $f'_x(x, y)$ и $f'_y(x, y)$, в свою очередь, будут функциями двух переменных x и y , определенными в той же области D или ее части. Будем называть их частными производными *первого порядка*.

Частные производные по x и по y от функций $f'_x(x, y)$ и $f'_y(x, y)$ в точке (x, y) , если они существуют, называются частными производными *второго порядка* от функции $z = f(x, y)$ в этой точке и обозначаются следующим образом:

$$\begin{aligned} f''_{x^2}(x, y) &= \frac{\partial^2 f(x, y)}{\partial x^2} = z''_{x^2} = \frac{\partial^2 z}{\partial x^2}, \\ f''_{xy}(x, y) &= \frac{\partial^2 f(x, y)}{\partial x \partial y} = z''_{xy} = \frac{\partial^2 z}{\partial x \partial y}, \\ f''_{yx}(x, y) &= \frac{\partial^2 f(x, y)}{\partial y \partial x} = z''_{yx} = \frac{\partial^2 z}{\partial y \partial x}, \\ f''_{y^2}(x, y) &= \frac{\partial^2 f(x, y)}{\partial y^2} = z''_{y^2} = \frac{\partial^2 z}{\partial y^2}. \end{aligned}$$

По определению

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} = \frac{\partial}{\partial y} \left(\frac{\partial f(x, y)}{\partial x} \right) \text{ и т. д.}$$

Частные производные второго порядка зависят от координат точки, в которой они вычисляются, т. е., в свою очередь, являются функциями двух переменных, определенными в области D или в ее части (если не во всех точках области D существуют эти частные производные).

Так, например, для функции $f(x, y) = x^4 + 3x^2y^2$ в любой точке (x, y) плоскости имеем:

$$\begin{aligned} f'_x(x, y) &= 4x^3 + 6xy^2, \quad f'_y(x, y) = 6x^2y, \\ f''_{x^2}(x, y) &= 12x^2 + 6y^2, \quad f''_{xy}(x, y) = 12xy, \quad f''_{yx}(x, y) = 12xy, \\ f''_{y^2}(x, y) &= 6x^2. \end{aligned}$$

В частности, $f''_{x^2}(1, 2) = 12 \cdot 1 + 6 \cdot 4 = 36$, $f''_{xy}(-1, 3) = 12 \cdot (-1) \cdot 3 = -36$ и т. д.

Частные производные *третьего, четвертого* и пр. *порядков* и соответствующая символика вводятся аналогично.

Частная производная любого порядка, взятая по различным переменным, как, например,

$$\frac{\partial^2 z}{\partial x \partial y}, \quad \frac{\partial^2 z}{\partial y \partial x}, \quad \frac{\partial^3 z}{\partial x \partial y \partial x}, \quad \frac{\partial^3 z}{\partial y \partial x \partial x}$$

и т. д., называется *смешанной*.

Очевидно, что для функции двух переменных можно, вообще говоря, определить две частные производные первого порядка, четыре частных производных второго порядка, восемь — третьего и вообще 2^m частных производных порядка m .

Для функции $f(x, y) = x^4 + 3x^2y^2$ (см. начатый выше пример) имеем:

$$\begin{aligned} f'''_{x^3}(x, y) &= 24x, \quad f'''_{x^2y}(x, y) = 12y, \quad f'''_{xyx}(x, y) = 12y, \\ f'''_{xy^2}(x, y) &= 12x, \quad f'''_{yx^2}(x, y) = 12y, \quad f'''_{yxy}(x, y) = 12x, \\ f'''_{y^3}(x, y) &= 12x, \quad f'''_{y^2x}(x, y) = 0 \text{ и т. д.} \end{aligned}$$

Сравнивая между собой значения смешанных производных функции $f(x, y) = x^4 + 3x^2y^2$ в любой точке (x, y) , видим, что

$$f''_{xy}(x, y) = f''_{yx}(x, y), \\ f'''_{x^2y}(x, y) = f'''_{xyx}(x, y) = f'''_{yx^2}(x, y), f'''_{xy^2}(x, y) = f'''_{yxy}(x, y) = f'''_{y^2x}(x, y),$$

т. е. смешанные частные производные данной функции, отличающиеся лишь последовательностью произведенных дифференцирований, совпадают друг с другом. Это подмеченное на конкретном примере свойство имеет место для широкого класса функций, что подтверждается следующей теоремой.

Теорема. Если в некоторой окрестности точки (x_0, y_0) функция $z = f(x, y)$ имеет смешанные частные производные $f''_{xy}(x, y)$ и $f''_{yx}(x, y)$ и эти производные непрерывны в самой точке (x_0, y_0) , то они в этой точке равны:

$$f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0).$$

Для доказательства рассмотрим вспомогательное выражение¹⁾:

$$A = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0) - f(x_0, y_0 + \Delta y) + f(x_0, y_0)^2).$$

Обозначив $\varphi(x) = f(x, y_0 + \Delta y) - f(x, y_0)$, видим, что

$$A = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)] - [f(x_0, y_0 + \Delta y) - f(x_0, y_0)] = \\ = \varphi(x_0 + \Delta x) - \varphi(x_0).$$

Применяем к этой разности теорему Лагранжа:

$$A = \varphi'(x_0 + \theta_1 \Delta x) \Delta x,$$

т. е.

$$A = [f'_x(x_0 + \theta_1 \Delta x, y_0 + \Delta y) - f'_x(x_0 + \theta_1 \Delta x, y_0)] \Delta x, \text{ где } 0 < \theta_1 < 1.$$

Справа в квадратных скобках стоит приращение функции $f'_x(x_0 + \theta_1 \Delta x, y)$ по y при изменении y от y_0 до $y_0 + \Delta y$. Применяя еще раз теорему Лагранжа, получаем:

$$A = f''_{xy}(x_0 + \theta_1 \Delta x, y_0 + \theta_2 \Delta y) \Delta x \Delta y, \text{ где } 0 < \theta_1 < 1, 0 < \theta_2 < 1.$$

Аналогичным образом, меняя ролями переменные x и y , находим:

$$A = f''_{yx}(x_0 + \theta_3 \Delta x, y_0 + \theta_4 \Delta y) \Delta x \Delta y, \text{ где } 0 < \theta_3 < 1, 0 < \theta_4 < 1,$$

откуда следует, что

$$f''_{xy}(x_0 + \theta_1 \Delta x, y_0 + \theta_2 \Delta y) = f''_{yx}(x_0 + \theta_3 \Delta x, y_0 + \theta_4 \Delta y).$$

При $\Delta x \rightarrow 0, \Delta y \rightarrow 0$, в силу непрерывности смешанных производных в точке (x_0, y_0) , получаем:

$$f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0).$$

Из доказанной теоремы вытекают два следствия:

1. Если смешанные производные второго порядка от функции $z = f(x, y)$ непрерывны в некоторой области, то они совпадают в этой области.

¹⁾ Предполагается, что точка $(x_0 + \Delta x, y_0 + \Delta y)$ лежит в данной окрестности точки (x_0, y_0) .

²⁾ Выражение A можно рассматривать как частное приращение по x (по y) от частного приращения по y (по x) от функции $z = f(x, y)$ в точке (x_0, y_0) , когда переменным x и y даются приращения Δx и Δy .

Например, $\Delta_y z = f(x_0, y_0 + \Delta y) - f(x_0, y_0)$,

$$\Delta_x (\Delta_y z) = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)] - \\ - [f(x_0, y_0 + \Delta y) - f(x_0, y_0)] = A.$$

Очевидно, что

$$\Delta_x (\Delta_y z) = \Delta_y (\Delta_x z).$$

2. Если функция $z=f(x, y)$ в некоторой области имеет все частные производные до n -го порядка включительно и эти производные непрерывны, то смешанные производные порядка m , $m \leq n$, отличающиеся лишь последовательностью дифференцирований, совпадают между собой.

Действительно, если имеются две любые смешанные производные порядка m , $m \leq n$, отличающиеся друг от друга только последовательностью дифференцирований, то от одной из них к другой можно перейти рядом перестановок двух соседних дифференцирований. А каждая такая перестановка оставляет, по доказанному выше, производную неизменной.

Например, если $n \geq 5$, то

$$f_{xxyxy}^{(5)}(x, y) = f_{xyxxx}^{(5)}(x, y).$$

В самом деле,

$$f_{xxy}''' = (f_x')_{xy}'' = (f_y')_{xx}'' = f_{xyx}''', \quad \text{т. е.} \quad f_{xxy}''' = f_{xyx}'''.$$

Дифференцируя обе части этого равенства сначала по x , потом по y , получаем:

$$f_{xxyxy}^{(5)}(x, y) = f_{xyxxx}^{(5)}(x, y).$$

Таким образом, значительно сокращается число различных между собой частных производных порядка m , $m \leq n$. Если последовательность дифференцирований для функции $f(x, y)$ безразлична, то любую частную производную порядка m можно получить, выполнив сначала все необходимые дифференцирования по x , а затем все положенные дифференцирования по y , т. е. любая частная производная порядка m такой функции может быть записана в виде:

$$\frac{\partial^m z}{\partial x^k \partial y^{m-k}},$$

где k — одно из чисел $0, 1, 2, \dots, m$.

Отсюда следует, что число различных частных производных порядка m у этой функции равно $m+1$, а не 2^m , как было указано раньше.

Примечание. Не следует думать, что смешанные производные любой функции не зависят от последовательности дифференцирований. Примеры показывают, что в случае, если смешанные производные не являются непрерывными, результат дифференцирования существенно зависит от последовательности дифференцирований.

Определения и обозначения частных производных высших порядков для функций трех и более переменных даются аналогичным образом. Остается в силе и теорема о независимости смешанных производных от последовательности дифференцирований при условии их непрерывности.

Например, если $u = x^3 y^2 z$, то в любой точке (x, y, z)

$$\begin{aligned} \frac{\partial u}{\partial x} &= 3x^2 y^2 z, & \frac{\partial u}{\partial y} &= 2x^3 y z, & \frac{\partial u}{\partial z} &= x^3 y^2, \\ \frac{\partial^2 u}{\partial x^2} &= 6x y^2 z, & \frac{\partial^2 u}{\partial x \partial y} &= \frac{\partial^2 u}{\partial y \partial x} = 6x^2 y z, & \frac{\partial^2 u}{\partial x \partial z} &= \frac{\partial^2 u}{\partial z \partial x} = 3x^2 y^2 \quad \text{и т. д.} \end{aligned}$$

§ 13. Дифференциалы высших порядков

Пусть $z=f(x, y)$ — дифференцируемая в области D функция двух независимых переменных x и y . В любой точке (x, y) этой области, давая x и y приращения $\Delta x = dx$ и $\Delta y = dy$, мы можем вычислить полный дифференциал:

$$dz = f'_x(x, y) dx + f'_y(x, y) dy.$$

Будем называть его дифференциалом *первого порядка*. Он зависит от значений x, y, dx и dy , т. е. является функцией четырех переменных.

Закрепляя dx и dy , получаем функцию двух переменных x и y , определенную в области D .

Дифференциал от этой функции в любой точке (x, y) области D , если он существует, называется дифференциалом *второго порядка* от функции $z=f(x, y)$ в точке (x, y) . Обозначается d^2z или $d^2f(x, y)$. Таким образом, $d^2z=d(dz)$.

Аналогично определяются дифференциалы *третьего, четвертого, ... , n -го порядков*. При этом

$$d^n z = d(d^{n-1}z).$$

Если функция $z=f(x, y)$ в области D имеет непрерывные частные производные до n -го порядка включительно, то и все дифференциалы, до n -го порядка включительно, в каждой точке области D существуют. Вычисляя их по известным правилам, получим:

$$\begin{aligned} dz &= f'_x(x, y) dx + f'_y(x, y) dy, \\ d^2z &= d[f'_x(x, y) dx + f'_y(x, y) dy] = df'_x(x, y) \cdot dx + df'_y(x, y) \cdot dy = \\ &= f''_{x^2}(x, y) (dx)^2 + 2f''_{xy}(x, y) dx dy + f''_{y^2}(x, y) (dy)^2, \\ d^3z &= d(d^2z) = d[f''_{x^2}(x, y) (dx)^2 + 2f''_{xy}(x, y) dx dy + f''_{y^2}(x, y) (dy)^2] = \\ &= f'''_{x^3}(x, y) (dx)^3 + 3f'''_{x^2y}(x, y) (dx)^2 dy + \\ &+ 3f'''_{xy^2}(x, y) dx (dy)^2 + f'''_{y^3}(x, y) (dy)^3 \text{ и т. д.} \end{aligned}$$

Назовем символической n -й степенью $\left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^n f(x, y)$ выражение, получающееся в результате возведения двучлена, стоящего в скобках, в обычную n -ю степень с последующей заменой степеней $\frac{\partial}{\partial x}$ и $\frac{\partial}{\partial y}$, умноженных на $f(x, y)$, частными производными соответствующих порядков от функции $f(x, y)$.

Тогда

$$\begin{aligned} dz &= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right) f(x, y), \\ d^2z &= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^2 f(x, y), \\ d^3z &= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^3 f(x, y) \text{ и т. д.} \end{aligned} \quad (1)$$

Можно показать, что $d^n z = \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^n f(x, y)$.

Определения и форма записи (1) дифференциалов высших порядков остаются в силе для функции любого числа независимых переменных.

Примечание. Если $z=f(x, y)$ — дифференцируемая функция промежуточных переменных x и y , являющихся, в свою очередь, дифференцируемыми функциями u и v , то, определяя $d^2z=d(dz)$, $d^3z=d(d^2z)$ и т. д., мы уже не получим указанных выше формул (1) для вычисления дифференциалов.

Так, например,

$$\begin{aligned} d^2z &= d[f'_x(x, y) dx + f'_y(x, y) dy] = df'_x(x, y) dx + \\ &+ f'_x(x, y) d(dx) + df'_y(x, y) dy + f'_y(x, y) d(dy) = \\ &= f''_{x^2}(x, y) (dx)^2 + 2f''_{xy}(x, y) dx dy + \\ &+ f''_{y^2}(x, y) (dy)^2 + f'_x(x, y) d^2x + f'_y(x, y) d^2y. \end{aligned}$$

Здесь $\Delta x \neq dx$, $\Delta y \neq dy$, dx и dy не являются постоянными (постоянны $du = \Delta u$ и $dv = \Delta v$).

Таким образом, указанная выше форма записи второго, третьего и т. д. дифференциалов инвариантной не является.

§ 14. Формула Тейлора для функции двух переменных

Пусть функция $z=f(x, y)$ имеет в окрестности точки (x_0, y_0) непрерывные частные производные до $(n+1)$ -го порядка включительно. Дадим x_0 и y_0 приращения Δx и Δy , не выводящие за пределы данной окрестности, и рассмотрим вспомогательную функцию

$$F(t) = f(x_0 + t\Delta x, y_0 + t\Delta y).$$

При постоянных $x_0, y_0, \Delta x$ и Δy это сложная функция переменной t , определенная в окрестности точки $t=0$ ¹⁾, имеющая в этой окрестности непрерывные производные до порядка $n+1$ включительно. Применяя к ней формулу Тейлора (см. гл. IV, § 15), получим:

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \dots + \frac{F^{(n)}(0)}{n!}t^n + \frac{F^{(n+1)}(\theta t)}{(n+1)!}t^{n+1}, \quad (1)$$

где $0 < \theta < 1$.

Так как $F(t) = f(x_0 + t\Delta x, y_0 + t\Delta y)$, то

$$\begin{aligned} F'(t) &= f'_x(x_0 + t\Delta x, y_0 + t\Delta y)\Delta x + f'_y(x_0 + t\Delta x, y_0 + t\Delta y)\Delta y = \\ &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right) f(x_0 + t\Delta x, y_0 + t\Delta y), \\ F''(t) &= f''_{x^2}(x_0 + t\Delta x, y_0 + t\Delta y)(\Delta x)^2 + \\ &+ 2f''_{xy}(x_0 + t\Delta x, y_0 + t\Delta y)\Delta x\Delta y + f''_{y^2}(x_0 + t\Delta x, y_0 + t\Delta y)(\Delta y)^2 = \\ &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^2 f(x_0 + t\Delta x, y_0 + t\Delta y), \\ &\dots\dots\dots \\ F^{(n)}(t) &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^n f(x_0 + t\Delta x, y_0 + t\Delta y), \\ F^{(n+1)}(t) &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^{n+1} f(x_0 + t\Delta x, y_0 + t\Delta y). \end{aligned}$$

Отсюда находим:

$$\begin{aligned} F(0) &= f(x_0, y_0), \\ F'(0) &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right) f(x_0, y_0), \\ F''(0) &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^2 f(x_0, y_0), \\ &\dots\dots\dots \\ F^{(n)}(0) &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^n f(x_0, y_0), \\ F^{(n+1)}(\theta t) &= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^{n+1} f(x_0 + \theta t\Delta x, y_0 + \theta t\Delta y), \quad 0 < \theta < 1. \end{aligned}$$

Найденные значения производных подставим в формулу (1), положив в ней предварительно $t=1$.

Учитывая, что $F(1) = f(x_0 + \Delta x, y_0 + \Delta y)$, получим:

$$\begin{aligned} f(x_0 + \Delta x, y_0 + \Delta y) &= f(x_0, y_0) + \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right) f(x_0, y_0) + \\ &+ \frac{1}{2!} \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^2 f(x_0, y_0) + \dots + \\ &+ \frac{1}{n!} \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^n f(x_0, y_0) + \\ &+ \frac{1}{(n+1)!} \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y \right)^{n+1} f(x_0 + \theta\Delta x, y_0 + \theta\Delta y), \quad 0 < \theta < 1. \end{aligned}$$

Эта формула называется *формулой Тейлора* для функции двух переменных. При $n=1$ она имеет вид:

$$\begin{aligned} f(x_0 + \Delta x, y_0 + \Delta y) &= f(x_0, y_0) + f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y + \\ &+ \frac{1}{2!} [f''_{x^2}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)(\Delta x)^2 + \\ &+ 2f''_{xy}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)\Delta x\Delta y + \\ &+ f''_{y^2}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)(\Delta y)^2], \quad 0 < \theta < 1. \end{aligned} \quad (2)$$

¹⁾ Радиус этой окрестности больше единицы.

Формула Тейлора удобно записывается в дифференциальной форме:

$$\begin{aligned}\Delta f &= f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \\ &= df(x_0, y_0) + \frac{1}{2!} d^2 f(x_0, y_0) + \dots + \frac{1}{n!} d^n f(x_0, y_0) + \\ &\quad + \frac{1}{(n+1)!} d^{n+1} f(x_0 + \theta \Delta x, y_0 + \theta \Delta y),\end{aligned}$$

где $0 < \theta < 1$.

Аналогичный вид имеет формула Тейлора для функций большего числа переменных.

Раздел 3

ЭКСТРЕМУМЫ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 15. Экстремумы функций нескольких переменных

Пусть $z = f(x, y)$ — функция двух переменных, определенная в некоторой области, (x_0, y_0) — точка этой области.

Точка (x_0, y_0) называется *точкой максимума* функции $f(x, y)$, если всюду в некоторой окрестности этой точки

$$f(x, y) \leq f(x_0, y_0)$$

(рис. 309).

Аналогично точка (x_0, y_0) называется *точкой минимума* функции $f(x, y)$, если всюду в некоторой окрестности этой точки

$$f(x, y) \geq f(x_0, y_0)$$

(рис. 310).

Если существует окрестность точки (x_0, y_0) , в которой $f(x, y) < f(x_0, y_0)$ ($f(x, y) > f(x_0, y_0)$), то точку (x_0, y_0) называют *точкой строгого максимума* (стро-

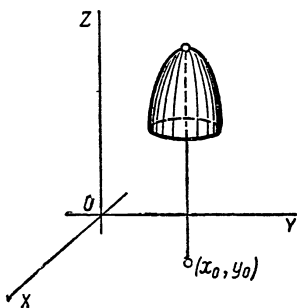


Рис. 309.

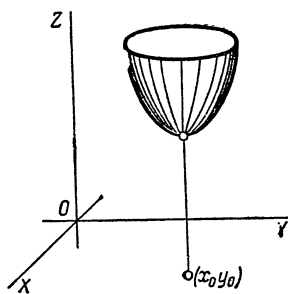


Рис. 310.

гого минимума) функции $f(x, y)$. В связи с этим определенные выше точки максимума и минимума называют иногда точками *нестромого* максимума и минимума.

В дальнейшем, говоря о точках максимума и минимума, мы будем иметь в виду данные выше определения.

Точки максимума и минимума функции называются ее *точками экстремума*.

Значения функции в точках максимума и минимума называются соответственно *максимумами* и *минимумами*, или, короче, *экстремумами*, этой функции.

Понятия экстремумов носят локальный характер: значение функции в точке (x_0, y_0) сравнивается со значениями функции в достаточно близких точках.

В данной области функция может совсем не иметь экстремумов, а может иметь несколько максимумов, несколько минимумов и даже бесчисленное множество и тех и других.

При этом некоторые минимумы функции могут оказаться больше некоторых из ее максимумов (рис. 311). Не следует смешивать максимумы и минимумы функции с ее наибольшими и наименьшими значениями.

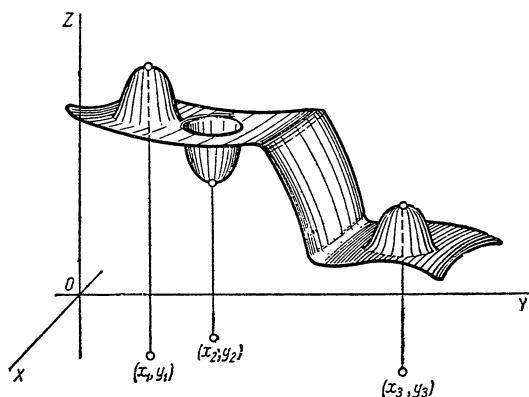


Рис. 311.

Покажем, что если в точке экстремума (x_0, y_0) функции $f(x, y)$ существуют частные производные этой функции: $f'_x(x_0, y_0)$ и $f'_y(x_0, y_0)$, то они равны нулю.

Пусть, например, (x_0, y_0) — точка максимума. Тогда по определению существует окрестность этой точки, для всех точек (x, y) которой

$$f(x, y) \leq f(x_0, y_0).$$

В частности, $f(x, y_0) \leq f(x_0, y_0)$ для всех x из некоторой окрестности $x_0 - \delta < x < x_0 + \delta$ точки x_0 . Это значит, что функция одной переменной $f(x, y_0)$, дифференцируемая в точке x_0 , имеет в ней максимум и, следовательно, в этой точке выполняется необходимое условие экстремума функции одной переменной:

$$f'_x(x_0, y_0) = 0.$$

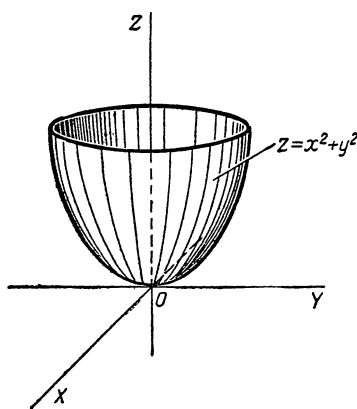


Рис. 312.

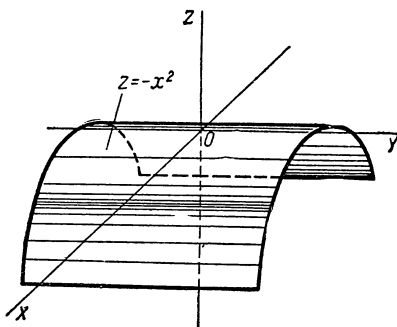


Рис. 313.

Аналогично, рассматривая функцию $f(x_0, y)$, находим:

$$f'_y(x_0, y_0) = 0.$$

Таким образом, мы получаем *необходимые условия экстремума* функции двух переменных.

Теорема. Если функция $f(x, y)$ в точке (x_0, y_0) имеет экстремум, то в этой точке либо обе ее частные производные первого порядка равны нулю: $f'_x(x_0, y_0) = 0$, $f'_y(x_0, y_0) = 0$, либо хотя бы одна из этих частных производных не существует.

Рассмотрим *примеры*.

1. Функция $f(x, y) = x^2 + y^2$ имеет минимум в точке $(0,0)$, так как $f(0,0) = 0$, $f(x, y) > 0$, если $x^2 + y^2 > 0$. При этом $f'_x(x, y) = 2x$, $f'_y(x, y) = 2y$, т. е. $f'_x(0,0) = f'_y(0,0) = 0$ (рис. 312).

2. Функция $f(x, y) = -x^2$ имеет максимум в любой точке оси OY (рис. 313). В любой точке $(0, y)$ оси OY

$$f'_x(0, y) = 0, \quad f'_y(0, y) = 0.$$

3. Функция $f(x, y) = 1 - \sqrt{x^2 + y^2}$ имеет максимум в точке $(0, 0)$, так как $f(0, 0) = 1$, $f(x, y) < 1$, если $x^2 + y^2 > 0$.
Частные производные функции $f(x, y)$:

$$f'_x(x, y) = \frac{-x}{\sqrt{x^2 + y^2}}, \quad f'_y(x, y) = \frac{-y}{\sqrt{x^2 + y^2}}$$

в точке $(0, 0)$ не существуют (рис. 314).

Пусть функция $f(x, y)$ задана в некоторой области. Будем называть ее *критическими точками* в этой области точки, в которых обе частные производные первого порядка одновременно равны нулю или хотя бы одна из этих производных не существует. Как было показано выше, экстремумы функции в данной области являются ее критическими точками. Однако не всякая критическая точка функции будет ее точкой экстремума.

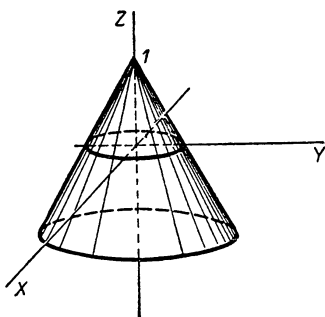


Рис. 314.

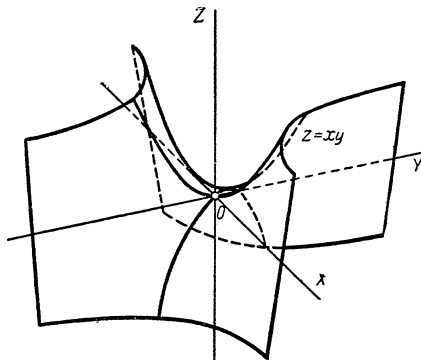


Рис. 315.

Например, функция $f(x, y) = xy$ имеет точку $(0, 0)$ критической точкой, так как $f'_x(0, 0) = f'_y(0, 0) = 0$.

При этом $f(0, 0) = 0$, а в любой сколь угодно малой окрестности точки $(0, 0)$ имеются точки, где $f(x, y) > 0$, и точки, где $f(x, y) < 0$, т. е. в точке $(0, 0)$ нет ни максимума, ни минимума функции. Поверхность, изображающая функцию $f(x, y) = xy$, есть гиперболический параболоид (рис. 315).

Таким образом, для нахождения экстремумов функции в данной области надо всякую критическую точку функции в этой области подвергнуть дополнительному исследованию.

Наиболее употребительные *достаточные условия экстремума* для функции двух переменных формулируются следующим образом.

Теорема. Пусть в окрестности критической точки (x_0, y_0) функция $f(x, y)$ имеет непрерывные частные производные до второго порядка включительно. Рассмотрим выражение

$$D(x, y) = f''_{x^2}(x, y) \cdot f''_{y^2}(x, y) - [f''_{xy}(x, y)]^2.$$

Тогда:

- 1) если $D(x_0, y_0) > 0$, то в точке (x_0, y_0) функция $f(x, y)$ имеет экстремум: максимум, если $f''_{x^2}(x_0, y_0) < 0$, и минимум, если $f''_{x^2}(x_0, y_0) > 0$;
- 2) если $D(x_0, y_0) < 0$, то в точке (x_0, y_0) функция $f(x, y)$ экстремума не имеет.

В случае, когда $D(x_0, y_0) = 0$, экстремум в точке (x_0, y_0) может быть, может и не быть.

Доказательство. Рассмотрим приращение Δz , получаемое функцией $z = f(x, y)$ при переходе от критической точки $M_0(x_0, y_0)$ к произвольной точке $M(x, y)$:

$$\Delta z = f(x, y) - f(x_0, y_0).$$

Чтобы выяснить, имеет ли функция $f(x, y)$ в точке M_0 экстремум, достаточно исследовать знак Δz . Если существует окрестность точки M_0 , для всех точек M которой

$\Delta z \geq 0$ ($\Delta z \leq 0$), то M_0 — точка минимума (максимума) функции $f(x, y)$. Если же в любой сколь угодно малой окрестности точки M_0 найдутся точки, в которых Δz как больше, так и меньше нуля, то в точке M_0 экстремума нет.

Для исследования знака Δz разложим функцию $f(x, y)$ в окрестности точки M_0 по формуле Тейлора, взяв $n = 1$ [(2), § 14]. Так как в точке M_0

$$f'_x(x_0, y_0) = f'_y(x_0, y_0) = 0,$$

то

$$\Delta z = \frac{1}{2} [f''_{xx}(x_0 + \theta \Delta x, y_0 + \theta \Delta y) (\Delta x)^2 + 2f''_{xy}(x_0 + \theta \Delta x, y_0 + \theta \Delta y) \Delta x \Delta y + \\ + f''_{yy}(x_0 + \theta \Delta x, y_0 + \theta \Delta y) (\Delta y)^2],$$

где $0 < \theta < 1$, $\Delta x = x - x_0$, $\Delta y = y - y_0$.

Обозначим через A, B, C соответственно числа $f''_{xx}(x_0, y_0)$, $f''_{xy}(x_0, y_0)$ и $f''_{yy}(x_0, y_0)$.

Тогда

$$\begin{aligned} f''_{xx}(x_0 + \theta \Delta x, y_0 + \theta \Delta y) &= A + \alpha, \\ f''_{xy}(x_0 + \theta \Delta x, y_0 + \theta \Delta y) &= B + \beta, \\ f''_{yy}(x_0 + \theta \Delta x, y_0 + \theta \Delta y) &= C + \gamma, \end{aligned}$$

причем, в силу непрерывности частных производных второго порядка в точке M_0 , при стремлении Δx и Δy к нулю α, β и γ также стремятся к нулю.

Вводя новые обозначения, получим:

$$\Delta z = \frac{1}{2} \{ [A (\Delta x)^2 + 2B \Delta x \Delta y + C (\Delta y)^2] + [\alpha (\Delta x)^2 + 2\beta \Delta x \Delta y + \gamma (\Delta y)^2] \}$$

или, полагая $\Delta x = \rho \cos \varphi$, $\Delta y = \rho \sin \varphi$ (здесь φ — угол, составляемый вектором «сдвига» M_0M с осью OX , а $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ — величина сдвига):

$$\begin{aligned} \Delta z = \frac{\rho^2}{2} [& (A \cos^2 \varphi + 2B \cos \varphi \sin \varphi + C \sin^2 \varphi) + \\ & + (\alpha \cos^2 \varphi + 2\beta \cos \varphi \sin \varphi + \gamma \sin^2 \varphi)]. \end{aligned} \quad (1)$$

Пользуясь этим представлением приращения Δz , покажем, что характер критической точки M_0 определяется знаком «дискриминанта» «квадратичного» трехчлена

$$A \cos^2 \varphi + 2B \cos \varphi \sin \varphi + C \sin^2 \varphi. \quad (2)$$

Рассмотрим возможные случаи.

1. $AC - B^2 > 0$.

Так как в этом случае $A \neq 0$, то трехчлен (2) можно записать в виде:

$$\frac{1}{A} [(A \cos \varphi + B \sin \varphi)^2 + (AC - B^2) \sin^2 \varphi]. \quad (3)$$

Выражение, стоящее в квадратных скобках, при всех значениях φ , $0 \leq \varphi \leq 2\pi$, положительно, как сумма двух неотрицательных слагаемых $(A \cos \varphi + B \sin \varphi)^2$ и $(AC - B^2) \sin^2 \varphi$, обращающихся в нуль при различных φ ¹⁾. Следовательно, трехчлен (2) при всех значениях φ имеет тот же знак, что и число A . Обозначим через m наименьшее значение абсолютной величины этого трехчлена на отрезке $[0, 2\pi]$. В силу непрерывности трехчлена (2) как функции φ на отрезке $[0, 2\pi]$ это значение m принимается при некотором φ , а значит, положительно:

$$|A \cos^2 \varphi + 2B \cos \varphi \sin \varphi + C \sin^2 \varphi| \geq m > 0, \quad 0 \leq \varphi \leq 2\pi.$$

С другой стороны, так как α, β и γ — бесконечно малые при $\rho \rightarrow 0$ функции ρ и φ , то из неравенства

$$|\alpha \cos^2 \varphi + 2\beta \cos \varphi \sin \varphi + \gamma \sin^2 \varphi| \leq |\alpha| + 2|\beta| + |\gamma|$$

следует, что трехчлен $\alpha \cos^2 \varphi + 2\beta \cos \varphi \sin \varphi + \gamma \sin^2 \varphi$ для достаточно малых ρ будет сколь угодно малым по абсолютной величине.

В частности, для достаточно малых ρ (а значит, $|\Delta x|$ и $|\Delta y|$) при всех значениях φ , $0 \leq \varphi \leq 2\pi$,

$$|\alpha \cos^2 \varphi + 2\beta \cos \varphi \sin \varphi + \gamma \sin^2 \varphi| < m.$$

Возвращаясь к представлению (1), заключаем, что существует окрестность точки M_0 , для всех точек M которой приращение Δz имеет один и тот же знак — знак трехчлена (2), совпадающий со знаком числа A .

¹⁾ $A \cos \varphi + B \sin \varphi = 0$ там, где $\operatorname{ctg} \varphi = -\frac{B}{A}$, а $(AC - B^2) \sin^2 \varphi = 0$ там, где $\sin \varphi = 0$, т. е. $\operatorname{ctg} \varphi$ не существует.

Отсюда следует, что в точке M_0 функция $f(x, y)$ имеет экстремум: максимум, если $A < 0$, и минимум, если $A > 0$.

2. $AC - B^2 < 0$.

В этом случае число A может быть как равно нулю, так и отлично от нуля.

Если $A = 0$, то трехчлен (2) запишется в виде:

$$2B \cos \varphi \sin \varphi + C \sin^2 \varphi = \sin \varphi [2B \cos \varphi + C \sin \varphi]. \quad (4)$$

Так как при этом $B \neq 0$, то можно найти столь малый положительный угол φ_1 , что

$$|C| \sin \varphi_1 < 2|B| \cos \varphi_1.$$

При $\varphi = \varphi_1$ знак суммы $2B \cos \varphi + C \sin \varphi$ совпадает со знаком первого члена. При $\varphi = \varphi_2 = 2\pi - \varphi_1$ знак первого члена и всей суммы остается прежним. А так как $\sin \varphi$ для углов φ_1 и φ_2 имеет взаимно противоположные знаки, то и трехчлен (4) для $\varphi = \varphi_1$ и $\varphi = \varphi_2$ имеет взаимно противоположные знаки. Поскольку трехчлен

$$\alpha \cos^2 \varphi + 2\beta \cos \varphi \sin \varphi + \gamma \sin^2 \varphi$$

при достаточно малых ρ сколь угодно мал по абсолютной величине как для $\varphi = \varphi_1$, так и для $\varphi = \varphi_2$, то знак Δz определится для этих φ знаком трехчлена (2). Следовательно, в любой достаточно малой окрестности точки M_0 найдутся точки M (например, точки, лежащие на лучах $\varphi = \varphi_1$ и $\varphi = \varphi_2$), для которых приращение Δz имеет взаимно противоположные знаки. В такой точке M_0 экстремума нет.

Пусть теперь $A \neq 0$. Запишем трехчлен (2) снова в виде (3):

$$\frac{1}{A} [(A \cos \varphi + B \sin \varphi)^2 + (AC - B^2) \sin^2 \varphi].$$

При $\varphi = 0$ этот трехчлен равен числу A . Для угла $\varphi = \varphi_3$, определенного из условия:

$$A \cos \varphi_3 + B \sin \varphi_3 = 0,$$

трехчлен (2) равен:

$$\frac{1}{A} (AC - B^2) \sin^2 \varphi_3,$$

т. е. имеет знак, противоположный знаку A .

Учитывая, что трехчлен $\alpha \cos^2 \varphi + 2\beta \cos \varphi \sin \varphi + \gamma \sin^2 \varphi$ при $\varphi = 0$ и $\varphi = \varphi_3$ для достаточно малых ρ сколь угодно мал по абсолютной величине, снова заключаем, что знаки Δz будут взаимно противоположными вдоль лучей $\varphi = 0$ и $\varphi = \varphi_3$ в достаточной близости к точке M_0 .

Следовательно, и в этом случае функция $f(x, y)$ в точке M_0 не имеет экстремума.

3. $AC - B^2 = 0$.

Если $AC - B^2 = 0$, то функция $f(x, y)$ в точке M_0 может как иметь, так и не иметь экстремума. Это подтверждается, например, рассмотрением функций $f(x, y) = x^4 + y^4$ и $f(x, y) = x^4 - y^4$ в окрестности точки $(0, 0)$.

Исследование приращения Δz , основанное на использовании формулы Тейлора при $n=1$, в этом случае, как правило, не приводит к окончательному результату, в связи с чем для исследования на экстремум функции $f(x, y)$ приходится привлекать частные производные более высоких порядков. Ввиду сложности этого вопроса в данном курсе он не рассматривается.

Рассмотрим пример на нахождение экстремумов функции.

Функция $f(x, y) = xy(1 - x - y)$ определена всюду. Ее частные производные: $f'_x(x, y) = y(1 - 2x - y)$, $f'_y(x, y) = x(1 - x - 2y)$ существуют всюду.

Для определения критических точек функции решаем систему уравнений:

$$\begin{cases} y(1 - 2x - y) = 0, \\ x(1 - x - 2y) = 0. \end{cases}$$

Получаем точки $M_1(0, 0)$, $M_2(0, 1)$, $M_3(1, 0)$, $M_4(\frac{1}{3}, \frac{1}{3})$.

Частные производные второго порядка

$$f''_{xx}(x, y) = -2y, \quad f''_{xy}(x, y) = 1 - 2x - 2y, \quad f''_{yy}(x, y) = -2x$$

непрерывны всюду,

$$D(x, y) = 4xy - (1 - 2x - 2y)^2.$$

Неравенства $D(0, 0) < 0$, $D(0, 1) < 0$, $D(1, 0) < 0$ показывают, что в точках M_1 , M_2 , M_3 экстремумов нет.

Вычисляя $D\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{1}{9} > 0$, заключаем, что в точке M_4 функция $f(x, y)$ имеет экстремум.

Так как $f''_{x^2}\left(\frac{1}{3}, \frac{1}{3}\right) < 0$, то M_4 — точка максимума,

$$z_{\max} = f\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{1}{9}\left(1 - \frac{1}{3} - \frac{1}{3}\right) = \frac{1}{27}.$$

Определения максимума и минимума и необходимые условия экстремума легко переносятся на случай функций трех и более переменных.

Пусть функция n переменных $u = f(x, y, z, \dots, w)$ определена в области, содержащей точку $(x_0, y_0, z_0, \dots, w_0)$. Точка $(x_0, y_0, z_0, \dots, w_0)$ называется *точкой максимума (минимума)* этой функции, если всюду в некоторой окрестности точки $(x_0, y_0, z_0, \dots, w_0)$

$$f(x, y, z, \dots, w) \leq f(x_0, y_0, z_0, \dots, w_0) \quad (f(x, y, z, \dots, w) \geq f(x_0, y_0, z_0, \dots, w_0)).$$

Если в точке максимума или минимума (короче, в *точке экстремума*) функция $f(x, y, z, \dots, w)$ имеет частные производные первого порядка, то все эти частные производные равны нулю.

Таким образом, если функция $f(x, y, z, \dots, w)$ в некоторой точке имеет экстремум, то либо в этой точке все ее частные производные первого порядка равны нулю, либо хотя бы одна из них не существует.

Достаточные условия экстремума для функций n переменных ($n \geq 3$) ввиду их сложности в данном курсе не рассматриваются.

§ 16. Наибольшее и наименьшее значения функции

Если функция $z = f(x, y)$ определена и непрерывна в замкнутой ограниченной области D , то она принимает в этой области свое наибольшее и наименьшее значения.

Пусть (x_1, y_1) — точка, в которой принимается наибольшее значение, а (x_2, y_2) — точка, в которой принимается наименьшее значение функции $f(x, y)$ в области D^1). Это значит, что для всех точек (x, y) области D

$$f(x, y) \leq f(x_1, y_1) \quad \text{и} \quad f(x, y) \geq f(x_2, y_2).$$

Точки (x_1, y_1) и (x_2, y_2) могут лежать как на границе области D , так и внутри нее.

Если точка (x_1, y_1) [также (x_2, y_2)] лежит внутри области D , то это точка максимума (минимума) функции $f(x, y)$, т. е. критическая точка функции $f(x, y)$ — внутри области D .

Поэтому для нахождения наибольшего и наименьшего значений функции $f(x, y)$ в области D нужно найти все ее критические точки внутри области D , вычислить значения функции в этих точках, затем найти наибольшее и

наименьшее значения функции на границе области D и из всех полученных таким образом значений выбрать наибольшее и наименьшее.

Пример. Найти наибольшее и наименьшее значения функции $f(x, y) = xy(1 - x - y)$ в замкнутой области, ограниченной линиями $x = 0$, $y + x = 2$, $y = 0$ (рис. 316).

Исследуя эту функцию на экстремум, мы нашли четыре критические точки $M_1(0, 0)$, $M_2(0, 1)$, $M_3(1, 0)$, $M_4\left(\frac{1}{3}, \frac{1}{3}\right)$, из которых лишь одна точка $M_4\left(\frac{1}{3}, \frac{1}{3}\right)$ лежит внутри данной области. Было установлено, что в этой точке функция имеет максимум, причем

$$z_{\max} = \frac{1}{27}.$$

¹⁾ Таких точек может быть несколько.

Исследуем значения функции на границе области. Пусть B — точка $(2, 0)$, а C — точка $(0, 2)$.

На участке OB имеем $y = 0$, т. е. $z \equiv 0$.

На участке CO имеем $x = 0$, т. е. $z \equiv 0$.

На участке BC имеем $y + x = 2$, т. е. $z = x(x - 2)$, причем $0 \leq x \leq 2$.

Вычисляя $\frac{dz}{dx} = 2x - 2$, получаем $\frac{dz}{dx} = 0$ при $x = 1$.

Так как $\frac{d^2z}{dx^2} = 2 > 0$, то $x = 1$ — точка минимума для функции $z = x(x - 2)$, $z_{\min} = -1$.

Учитывая, что $z_B = 0$, $z_C = 0$, заключаем: на участке BC наименьшее значение функции $z = xy(1 - x - y)$ равно -1 , а наибольшее равно нулю.

Сравнивая найденные значения, приходим к выводу, что наибольшее значение функции $z = xy(1 - x - y)$ в данной области достигается внутри области, в точке $M_4\left(\frac{1}{3}, \frac{1}{3}\right)$; оно равно $\frac{1}{27}$. Наименьшее значение достигается на границе области, в точке $(1, 1)$; оно равно -1 .

§ 17. Условный экстремум

Выше ставилась задача об отыскании максимумов и минимумов функций нескольких переменных. При этом на изменение независимых переменных не накладывалось никаких ограничений.

На практике часто приходится иметь дело с исследованием на экстремум функций нескольких переменных при наличии определенных условий на эти переменные.

Например, рассматриваются такие задачи: на поверхности, заданной уравнением $F(x, y, z) = 0$, найти точки, в которых данная функция $f(x, y, z)$ принимает наибольшие и наименьшие значения по сравнению с ее значениями в близлежащих точках той же поверхности, или на кривой, заданной уравнениями $F_1(x, y, z) = 0$, $F_2(x, y, z) = 0$, найти точки, в которых данная функция $f(x, y, z)$ принимает наибольшие или наименьшие значения по сравнению с ее значениями в близлежащих точках той же кривой.

В обеих этих задачах экстремальные значения функции $f(x, y, z)$ ищутся среди точек, координаты которых x, y и z удовлетворяют одному ($F(x, y, z) = 0$) или двум ($F_1(x, y, z) = 0$, $F_2(x, y, z) = 0$) условиям, в связи с чем такие экстремумы и называются *условными*.

В общем виде задача на условный экстремум может быть поставлена следующим образом.

Пусть в некоторой области заданы функция n переменных $u = f(x, y, z, \dots, \omega)$ и m уравнений ($m < n$), связывающих эти переменные:

$$F_1(x, y, z, \dots, \omega) = 0, \quad F_2(x, y, z, \dots, \omega) = 0, \dots, \quad F_m(x, y, z, \dots, \omega) = 0.$$

Эти уравнения называются *уравнениями связи*. Будем говорить, что в точке $(x_0, y_0, z_0, \dots, \omega_0)$, координаты которой удовлетворяют уравнениям связи, функция $u = f(x, y, z, \dots, \omega)$ имеет *условный максимум (минимум)*, если неравенства

$$f(x, y, z, \dots, \omega) \leq f(x_0, y_0, z_0, \dots, \omega_0) \quad (f(x, y, z, \dots, \omega) \geq f(x_0, y_0, z_0, \dots, \omega_0))$$

выполняются для всех точек $(x, y, z, \dots, \omega)$ некоторой окрестности точки $(x_0, y_0, z_0, \dots, \omega_0)$, координаты которых также удовлетворяют уравнениям связи.

Решить задачу на условный экстремум для функции $u = f(x, y, z, \dots, \omega)$ при наличии указанных уравнений связи — значит найти все ее условные максимумы и минимумы в данной области.

Рассмотрим подробнее задачу на экстремум для функции $u = f(x, y, z)$ трех переменных, подчиненных двум уравнениям связи

$$F_1(x, y, z) = 0, \quad F_2(x, y, z) = 0. \quad (1)$$

Здесь $n = 3$, $m = 2$. Все рассуждения, проводимые нами в дальнейшем, будут справедливы для любых n и m ($n > m$).

Выведем **необходимые условия для условного экстремума**.

Пусть в точке $M_0(x_0, y_0, z_0)$, координаты которой удовлетворяют уравнениям (1), функция $u = f(x, y, z)$ имеет условный экстремум.

Предположим, что функции $f(x, y, z)$, $F_1(x, y, z)$ и $F_2(x, y, z)$ в окрестности точки (x_0, y_0, z_0) имеют непрерывные частные производные первого порядка по всем аргументам и хотя бы один из определителей второго порядка, составленный из элементов следующей таблицы:

$$\begin{vmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{vmatrix},$$

например,

$$\frac{\partial (F_1, F_2)}{\partial (y, z)} = \begin{vmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{vmatrix}$$

в точке M_0 отличен от нуля.

Тогда в некоторой окрестности точки M_0 на уравнения связи можно смотреть как на неявное задание однозначных функций $y = \varphi(x)$, $z = \psi(x)$ (непрерывных, имеющих непрерывные производные в окрестности x_0 и таких, что $\varphi(x_0) = y_0$, $\psi(x_0) = z_0$) и условный экстремум функции $f(x, y, z)$ в точке M_0 будет совпадать с безусловным экстремумом сложной функции $u = f(x, \varphi(x), \psi(x))$ в точке x_0 .

Отсюда следует, что если, разрешив систему (1) относительно переменных y и z , мы можем фактически найти явные выражения для функций $y = \varphi(x)$ и $z = \psi(x)$, то для отыскания точки $M_0(x_0, y_0, z_0)$ условного экстремума функции $u = f(x, y, z)$ нам достаточно решить задачу на безусловный экстремум для функции $u = f(x, \varphi(x), \psi(x))$.

Пример. Найти экстремумы функции $u = x^2 + y^2 + z^2$, если x, y и z удовлетворяют уравнениям $x + y - 3z + 7 = 0$, $x - y + z - 3 = 0$.

Разрешая уравнения связи относительно y и z и подставляя найденные значения $y = 2x - 1$, $z = x + 2$ под знак функции u , получаем функцию одной переменной x :

$$u = x^2 + (2x - 1)^2 + (x + 2)^2.$$

Экстремумы этой функции совпадают с условными экстремумами данной функции при наличии указанных уравнений связи.

Исследуем функцию $u = x^2 + (2x - 1)^2 + (x + 2)^2$ на экстремум: $u' = 2x + 4(2x - 1) + 2(x + 2)$, $u' = 0$ при $x_0 = 0$; так как $u' < 0$ при $x < 0$ и $u' > 0$ при $x > 0$, заключаем, что при $x_0 = 0$ функция имеет минимум, причем $u_{\min} = 5$.

Определив $y_0 = 2 \cdot 0 - 1 = -1$, $z_0 = 0 + 2 = 2$, мы можем сказать, что функция $u = x^2 + y^2 + z^2$ имеет единственный условный экстремум в точке $M_0(0, -1, 2)$: $u_{\min} = 5$.

Геометрически полученный результат значит, что среди всех точек $M(x, y, z)$ прямой, заданной уравнениями $x + y - 3z + 7 = 0$, $x - y + z - 3 = 0$, ближайшей к началу координат (u — квадрат расстояния точки от начала координат) будет точка $M_0(0, -1, 2)$.

Возвращаясь к поставленной выше общей задаче на экстремум функции $u = f(x, y, z)$ при наличии уравнений связи (1), укажем способ нахождения точки M_0 , не требующий фактического выражения переменных y и z через x из уравнений связи. При этом мы будем пользоваться только фактом существования неявных функций $y = \varphi(x)$ и $z = \psi(x)$.

Итак, пусть в точке $M_0(x_0, y_0, z_0)$ функция $u = f(x, y, z)$ имеет условный экстремум или, что то же, в точке x_0 дифференцируемая сложная функция $u = f(x, \varphi(x), \psi(x))$ имеет безусловный экстремум.

В такой точке x_0 (см. необходимое условие экстремума для функции одной переменной) $\frac{du}{dx} = 0$, откуда следует, что координаты точки $M_0(x_0, y_0, z_0)$ удовлетворяют соотношению

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx} + \frac{\partial f}{\partial z} \frac{dz}{dx} = 0. \quad (2)$$

Далее, так как в окрестности точки x_0 имеют место тождества:

$$F_1(x, \varphi(x), \psi(x)) \equiv 0, \quad F_2(x, \varphi(x), \psi(x)) \equiv 0,$$

то, дифференцируя их по x , мы получим соотношения, которым также удовлетворяют координаты точки M_0 :

$$\frac{\partial F_1}{\partial x} + \frac{\partial F_1}{\partial y} \frac{dy}{dx} + \frac{\partial F_1}{\partial z} \frac{dz}{dx} = 0, \quad (3)$$

$$\frac{\partial F_2}{\partial x} + \frac{\partial F_2}{\partial y} \frac{dy}{dx} + \frac{\partial F_2}{\partial z} \frac{dz}{dx} = 0. \quad (4)$$

Из уравнений (3) и (4) мы можем выразить производные $\frac{dy}{dx}$ и $\frac{dz}{dx}$ через x , y и z ¹⁾.

Соотношение (2), в котором $\frac{dy}{dx}$ и $\frac{dz}{dx}$ заменены их выражениями через x, y и z , и два уравнения связи (1) представляют собой систему из трех уравнений с тремя

¹⁾ Из условия следует, что определитель из коэффициентов при $\frac{dy}{dx}$ и $\frac{dz}{dx}$ в окрестности точки x_0 отличен от нуля.

неизвестными x, y, z , которым, как мы показали, удовлетворяют координаты точки M_0 условного экстремума.

Будем называть *критическими точками* для данной задачи точки (x, y, z) , координаты которых удовлетворяют указанной системе из трех уравнений.

Таким образом мы получим необходимые условия для условного экстремума: если функции f, F_1, F_2 удовлетворяют указанным выше условиям, то всякая точка M_0 условного экстремума есть критическая точка для данной задачи.

Будет ли каждая критическая точка точкой условного экстремума и, если будет, то каков характер этого экстремума, — выясняется дополнительно, исходя из конкретного смысла задачи или каких-либо еще соображений.

Достаточных условий для условного экстремума ввиду их сложности мы не рассматриваем.

На практике для определения координат точек условного экстремума обычно применяется *метод неопределенных множителей Лагранжа*, к изложению которого мы переходим.

Умножим соотношения (3) и (4) на произвольные пока числа (неопределенные множители) λ_1 и λ_2 и сложим с соотношениями (2). При этом получим равенство

$$\left(\frac{\partial f}{\partial x} + \lambda_1 \frac{\partial F_1}{\partial x} + \lambda_2 \frac{\partial F_2}{\partial x}\right) + \left(\frac{\partial f}{\partial y} + \lambda_1 \frac{\partial F_1}{\partial y} + \lambda_2 \frac{\partial F_2}{\partial y}\right) \frac{dy}{dx} + \\ + \left(\frac{\partial f}{\partial z} + \lambda_1 \frac{\partial F_1}{\partial z} + \lambda_2 \frac{\partial F_2}{\partial z}\right) \frac{dz}{dx} = 0,$$

справедливое в точке $M_0(x_0, y_0, z_0)$ условного экстремума.

Выберем теперь числа λ_1 и λ_2 так, чтобы вторая и третья скобки в точке M_0 были равны нулю. (Это можно сделать, и притом единственным образом, так как определитель $\frac{\partial(F_1, F_2)}{\partial(y, z)}$ в точке M_0 отличен от нуля.)

Тогда первая скобка также будет равна нулю, т. е. в точке M_0 будут выполнены соотношения:

$$\begin{aligned} \frac{\partial f}{\partial x} + \lambda_1 \frac{\partial F_1}{\partial x} + \lambda_2 \frac{\partial F_2}{\partial x} &= 0, \\ \frac{\partial f}{\partial y} + \lambda_1 \frac{\partial F_1}{\partial y} + \lambda_2 \frac{\partial F_2}{\partial y} &= 0, \\ \frac{\partial f}{\partial z} + \lambda_1 \frac{\partial F_1}{\partial z} + \lambda_2 \frac{\partial F_2}{\partial z} &= 0. \end{aligned}$$

Добавляя к ним уравнения связи $F_1=0, F_2=0$, получаем систему из пяти уравнений с пятью неизвестными $x, y, z, \lambda_1, \lambda_2$, которой при определенных значениях λ_1, λ_2 удовлетворяют координаты точки $M_0(x_0, y_0, z_0)$ условного экстремума.

В полученные по методу Лагранжа уравнения переменные x, y и z входят симметричным образом, что облегчает их запоминание.

Рассмотрим теперь вспомогательную функцию пяти переменных:

$$\Phi(x, y, z, \lambda_1, \lambda_2) = f(x, y, z) + \lambda_1 F_1(x, y, z) + \lambda_2 F_2(x, y, z),$$

где λ_1 и λ_2 — неопределенные числовые множители.

Полученная выше система из пяти уравнений может быть записана в виде:

$$\frac{\partial \Phi}{\partial x} = 0, \quad \frac{\partial \Phi}{\partial y} = 0, \quad \frac{\partial \Phi}{\partial z} = 0, \quad \frac{\partial \Phi}{\partial \lambda_1} = 0, \quad \frac{\partial \Phi}{\partial \lambda_2} = 0.$$

Так как это — необходимые условия безусловного экстремума функции $\Phi(x, y, z, \lambda_1, \lambda_2)$, заключаем, что всякая точка $M_0(x_0, y_0, z_0)$ условного экстремума функции $f(x, y, z)$ дает при специально подобранных λ_1 и λ_2 критическую точку безусловного экстремума функции $\Phi(x, y, z, \lambda_1, \lambda_2)$.

Аналогичным образом, решая задачу на экстремум функции n переменных

$$u = f(x, y, z, \dots, \omega)$$

при наличии m ($m < n$) уравнений связи

$$F_1 = 0, \quad F_2 = 0, \quad \dots, \quad F_m = 0, \quad (5)$$

при определенных условиях на функции $f, F_1, F_2, \dots, F_m$ получаем **необходимые условия** для условного экстремума: если в точке $(x_0, y_0, z_0, \dots, \omega_0)$ функция $u = f(x, y, z, \dots, \omega)$ имеет условный экстремум, то точка $(x_0, y_0, z_0, \dots, \omega_0, \lambda_1, \dots, \lambda_m)$ при некоторых значениях $\lambda_1, \lambda_2, \dots, \lambda_m$ будет критической точкой безусловного экстремума функции

$$\Phi = f + \lambda_1 F_1 + \lambda_2 F_2 + \dots + \lambda_m F_m.$$

Отсюда следует практическое правило для решения задачи на условный экстремум функции $u = f(x, y, z, \dots, \omega)$ при наличии уравнений связи (5).

1. Составляем вспомогательную функцию:

$$\Phi = f + \lambda_1 F_1 + \lambda_2 F_2 + \dots + \lambda_m F_m.$$

2. Записываем необходимые условия экстремума такой функции:

$$\frac{\partial \Phi}{\partial x} = \frac{\partial \Phi}{\partial y} = \frac{\partial \Phi}{\partial z} = \dots = \frac{\partial \Phi}{\partial w} = \frac{\partial \Phi}{\partial \lambda_1} = \dots = \frac{\partial \Phi}{\partial \lambda_m} = 0.$$

3. Решая полученную систему уравнений, находим критические точки экстремума для функции Φ .

4. Если $(x_0, y_0, z_0, \dots, w_0, \lambda_1, \dots, \lambda_m)$ — найденная критическая точка, то в точке $(x_0, y_0, z_0, \dots, w_0)$ функция $u = f(x, y, z, \dots, w)$, может быть, имеет условный экстремум. Это надо исследовать дополнительно.

Пример. Найти прямоугольный параллелепипед наибольшего объема, если его полная поверхность имеет площадь $2a$.

Пусть длины сторон параллелепипеда равны x , y и z . Его объем $V = xyz$, а площадь поверхности $S = 2(xy + xz + yz)$. Следовательно, требуется найти максимум функции $V = xyz$ при условии $xy + xz + yz - a = 0$.

Составляем вспомогательную функцию:

$$\Phi(x, y, z, \lambda) = xyz + \lambda(xy + xz + yz - a)$$

и приравниваем к нулю все ее частные производные первого порядка:

$$yz + \lambda(y + z) = 0, \quad (1')$$

$$xz + \lambda(x + z) = 0, \quad (2')$$

$$xy + \lambda(x + y) = 0, \quad (3')$$

$$xy + xz + yz - a = 0. \quad (4')$$

Учитывая, что x , y и z по смыслу задачи положительны, из соотношений (1') и (2') получаем:

$$\frac{yz}{y+z} = \frac{xz}{x+z},$$

т. е.

$$yz(x+z) = xz(y+z), \quad z(x-y) = 0, \quad x=y.$$

Аналогично из соотношений (2') и (3') вытекает, что $y=z$.

Подставляя в соотношение (4')

$$x=y=z,$$

находим:

$$3x^3 = a, \quad x = \sqrt[3]{\frac{a}{3}} = y = z^1).$$

Точка $\left(x_0 = \sqrt[3]{\frac{a}{3}}, y_0 = \sqrt[3]{\frac{a}{3}}, z_0 = \sqrt[3]{\frac{a}{3}}\right)$ является критической точкой экстремума функции V . Так как поставленная задача имеет определенное решение, а критическая точка только одна, то в этой критической точке будет экстремум.

Следовательно, искомый параллелепипед — куб со стороной $\sqrt[3]{\frac{a}{3}}$, его объем $V = \frac{a}{3} \sqrt[3]{\frac{a}{3}}$.

¹⁾ Соответствующее значение $\lambda = \frac{1}{2} \sqrt[3]{\frac{a}{3}}$ для нахождения критических точек экстремума функции V определять не требуется.

ГЛАВА VIII

ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Раздел I

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ПЕРВОГО ПОРЯДКА

§ 1. Задачи, приводящие к дифференциальным уравнениям. Основные определения

При решении многих геометрических и физических задач приходится отыскивать неизвестную функцию по данному соотношению между этой неизвестной функцией, ее производными и независимыми переменными. Такое соотношение называется *дифференциальным уравнением*, а отыскание функции, удовлетворяющей уравнению, называется *решением*, или *интегрированием*, данного уравнения.

Рассмотрим несколько задач, приводящих к дифференциальным уравнениям.

Задача 1. Найти кривую, обладающую тем свойством, что отрезок любой ее касательной, заключенный между осями координат, делится пополам в точке касания.

Пусть $y=f(x)$ — уравнение искомой кривой, $M(x, y)$ — произвольная точка кривой (рис. 317).

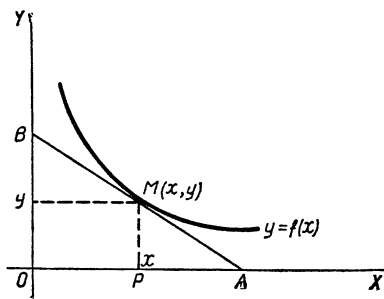


Рис. 317.

Угловой коэффициент касательной в этой точке равен y' . По условию $AM=MB$, т. е. $\overline{OP}=\overline{PA}=x$, а значит, в любой точке M кривой $\operatorname{tg} \angle MAP = -y' = \frac{y}{x}$; следовательно, $y' = -\frac{y}{x}$.

Мы получили соотношение, связывающее неизвестную функцию y , независимую переменную x и производную от y по x , т. е. получили дифференциальное уравнение относительно y . Этому уравнению удовлетворяет функция $y = \frac{C}{x}$, где C — любое число¹⁾.

Действительно, если $y = \frac{C}{x}$, то $y' = -\frac{C}{x^2}$ и $-\frac{y}{x} = -\frac{C}{x^2}$.

¹⁾ Методы для определения решений некоторых типов дифференциальных уравнений (и в частности, данного) будут указаны ниже.

Таким образом, указанным выше свойством обладает бесчисленное множество кривых («семейство» кривых), различающихся значениями постоянной C . Это — семейство равносторонних гипербол, асимптотами которых являются оси координат.

Для того чтобы из этого семейства кривых выделить одну определенную кривую, достаточно задать точку (x_0, y_0) , через которую проходит эта кривая, и определить соответствующее значение постоянной C .

Например, через точку $(2, 4)$ будет проходить та кривая семейства, для которой $4 = \frac{C}{2}$, т. е. $C = 8$. Уравнение этой кривой $y = \frac{8}{x}$.

Задача 2. (задача о радиоактивном распаде). Экспериментальным путем установлено, что скорость радиоактивного распада в каждый момент времени пропорциональна наличному количеству радиоактивного вещества. При этом предполагается, что количество радиоактивного вещества в породе настолько мало, что не вызывает цепной реакции. Требуется найти закон распада вещества, т. е. найти зависимость количества радиоактивного вещества в породе от времени t .

Решение. Пусть m — количество радиоактивного вещества в породе в момент времени t . Скорость изменения количества вещества равна $\frac{dm}{dt}$. Обозначив через k положительный коэффициент пропорциональности, запишем основной закон радиоактивного распада в виде:

$$\frac{dm}{dt} = -km$$

(знак минус берется потому, что скорость распада $\frac{dm}{dt}$ отрицательна).

Полученное соотношение представляет собой дифференциальное уравнение, связывающее искомую функцию m с ее производной $\frac{dm}{dt}$.

Легко убедиться в том, что любая функция

$$m = Ce^{-kt},$$

где C — число, удовлетворяет этому уравнению.

Действительно, $\frac{dm}{dt} = -kCe^{-kt}$, $-km = -kCe^{-kt}$,

т. е.

$$\frac{dm}{dt} \equiv -km.$$

Так как C — произвольное число, уравнение имеет бесчисленное множество решений, различающихся значениями постоянной C . Для того чтобы задача стала конкретной, достаточно указать количество радиоактивного вещества в породе в некоторый («начальный») момент времени t_0 . Пусть при $t = t_0$ в породе имелось m_0 граммов вещества. Тогда из соотношения $m_0 = C_0 e^{-kt_0}$ определится соответствующая решению постоянная $C_0 = m_0 e^{kt_0}$ и мы получим закон распада в виде:

$$m = m_0 e^{-k(t-t_0)}.$$

С помощью этого соотношения можно определить период полураспада вещества T , т. е. время, в течение которого количество вещества уменьшится вдвое.

Для этого положим $m = \frac{m_0}{2}$, а $t - t_0 = T$.

$$\frac{m_0}{2} = m_0 e^{-kT}, \quad e^{-kT} = \frac{1}{2},$$

$$T = \frac{\ln 2}{k} \approx \frac{0,693}{k}.$$

Постоянная k (коэффициент радиоактивности вещества) предполагается известной.

Если коэффициент радиоактивности k не известен, но известен период полураспада вещества, то $k = \frac{\ln 2}{T}$ и закон распада запишется в виде:

$$m = m_0 \left(\frac{1}{2} \right)^{\frac{t-t_0}{T}}.$$

Задача 3. Материальная точка массы m движется прямолинейно, притягиваемая к неподвижному центру O силой, пропорциональной расстоянию от точки до этого центра. Найти закон движения точки.

Примем точку O за начало, а прямую, по которой движется точка, за ось OY (рис. 318).

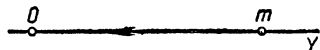


Рис. 318.

Притягивающая сила $P = -ky$, где y — координата точки m , а k — коэффициент пропорциональности ($k > 0$). Так как по второму закону Ньютона сила равна произведению массы m на ускорение $\frac{d^2y}{dt^2}$, мы получаем дифференциальное уравнение движения точки:

$$m \frac{d^2y}{dt^2} = -ky.$$

Это уравнение связывает искомую функцию y , ее вторую производную $\frac{d^2y}{dt^2}$ и независимую переменную — время t .

Легко проверить, что функция

$$y = C_1 \cos \sqrt{\frac{k}{m}} t + C_2 \sin \sqrt{\frac{k}{m}} t$$

при любых значениях чисел C_1 и C_2 удовлетворяет уравнению. Постоянные C_1 и C_2 , от которых зависит решение уравнения, определяются, если будут заданы конкретные «начальные» условия движения.

Пусть, например, известно, что в момент времени $t = 0$ материальная точка находилась на расстоянии a_0 от точки O и имела скорость v_0 .

Тогда, подставляя в соотношения

$$y = C_1 \cos \omega t + C_2 \sin \omega t, \quad \omega = \sqrt{\frac{k}{m}},$$

$$y' = -C_1 \omega \sin \omega t + C_2 \omega \cos \omega t$$

вместо t значение $t = 0$, получим $a_0 = C_1$, $v_0 = \omega C_2$, т. е.

$$C_1 = a_0, \quad C_2 = \frac{v_0}{\omega}.$$

Таким образом, искомый закон движения имеет вид:

$$y = a_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t.$$

Полагая $a_0 = R \sin \alpha$, $\frac{v_0}{\omega} = R \cos \alpha$, мы можем записать закон движения в виде:

$$y = R \sin (\alpha + \omega t)$$

(здесь $R = \sqrt{a_0^2 + \frac{v_0^2}{\omega^2}}$, $\operatorname{tg} \alpha = \frac{a_0 \omega}{v_0}$), откуда заключаем, что рассматриваемое движение есть периодическое колебательное движение, R — амплитуда, α — начальная фаза, $\omega = \sqrt{\frac{k}{m}}$ — частота колебания.

Дадим теперь точные определения основных понятий.

Обыкновенным дифференциальным уравнением n -го порядка называется соотношение вида:

$$F(x, y, y', \dots, y^{(n)}) = 0,$$

где F — функция, определенная в некоторой области, x — независимая переменная, y — искомая функция переменной x , а $y', y'', \dots, y^{(n)}$ — ее производные.

При этом функция F может явно не зависеть от $x, y, y', \dots, y^{(n-1)}$, но обязательно должна зависеть от $y^{(n)}$ ¹⁾.

Порядком дифференциального уравнения называется порядок старшей входящей в него производной.

Так, например, уравнения

$$y' + xy - x^2 = 0, \quad xy'^2 + e^x = 0, \quad yy' - 1 = 0, \quad y'^5 + e^{y^2} = 0$$

будут дифференциальными уравнениями первого порядка, уравнения $xy'' - y'^3 - y = 0$, $y'' - y' = 1$ будут дифференциальными уравнениями второго порядка, уравнение $y^2 - y''' + x^5 = 0$ имеет третий порядок и т. д.

Решением дифференциального уравнения называется всякая функция $y = \varphi(x)$, подстановка которой в это уравнение обращает его в тождество.

Например, дифференциальное уравнение $y'' + y = 0$ имеет решением функцию $y = \cos x$, так как если $y = \cos x$, то $y' = -\sin x$, $y'' = -\cos x$ и $-\cos x + \cos x \equiv 0$.

Решение дифференциального уравнения, заданное неявно соотношением $\Phi(x, y) = 0$, называют *интегралом* этого уравнения.

График решения дифференциального уравнения называется его *интегральной кривой*.

§ 2. Дифференциальные уравнения первого порядка, разрешенные относительно производной, и их геометрический смысл.

Основные понятия

Как уже было сказано, *дифференциальным уравнением первого порядка* называется соотношение между независимым переменным, неизвестной функцией и ее производной.

Общий вид дифференциального уравнения первого порядка:

$$F(x, y, y') = 0. \quad (1)$$

Если соотношение (1) разрешить относительно y' , то получится уравнение вида:

$$y' = f(x, y). \quad (2)$$

Оно называется дифференциальным уравнением первого порядка, *разрешенным относительно производной*.

Дифференциальное уравнение первого порядка, разрешенное относительно производной, всегда можно записать в так называемой *дифференциальной форме*:

$$P(x, y) dx + Q(x, y) dy = 0. \quad (3)$$

Действительно, если $y' = f(x, y)$, то $\frac{dy}{dx} = f(x, y)$, а значит,

$$f(x, y) dx - dy = 0.$$

Наоборот, всякое уравнение вида (3), если $Q(x, y) \neq 0$, можно разрешить относительно производной:

$$\frac{dy}{dx} = -\frac{P(x, y)}{Q(x, y)},$$

т. е. записать в виде: $y' = f(x, y)$.

¹⁾ В обыкновенных дифференциальных уравнениях искомая функция зависит только от одного аргумента, в отличие от *уравнений в частных производных*, в которых неизвестные функции зависят от нескольких независимых переменных. В дальнейшем, говоря о дифференциальных уравнениях, мы будем иметь в виду только обыкновенные дифференциальные уравнения.

Выясним геометрический смысл уравнения (2).

Пусть x и y — декартовы прямоугольные координаты точек плоскости, а $y = \varphi(x)$ — решение данного уравнения. График этого решения — интегральная кривая уравнения (2) — есть непрерывная кривая, в каждой точке которой имеется касательная. Угловым коэффициентом касательной к интегральной кривой в ее точке (x, y) равен y' , т. е. равен $f(x, y)$. Таким образом, уравнение $y' = f(x, y)$ дает связь между координатами точки и угловым коэффициентом касательной к интегральной кривой в этой точке.

В каждой точке (x, y) области D , в которой определена функция $f(x, y)$, мы можем вычислить y' , т. е. указать направление касательной к той интегральной кривой, которая проходит через эту точку¹⁾. Построив в каждой точке области черточку («стрелку»), наклоненную к оси OX под углом, тангенс которого равен $f(x, y)$, получим так называемое «поле направлений» (рис. 319).

Задать уравнение $y' = f(x, y)$ — значит задать в области D поле направлений²⁾. Решить это уравнение — значит найти кривую, касательная к которой в каждой ее точке совпадала бы с направлением поля в этой точке.

Таких кривых будет не одна, а целое семейство (построение можно начинать с любой точки данной области). Чтобы выделить определенную интегральную кривую, нужно задать точку (x_0, y_0) , через которую должна проходить кривая. При некоторых ограничениях на правую часть уравнения (2) через каждую точку области D будет проходить одна интегральная кривая.

Пример. Рассмотрим уравнение $y' = x + y$.

Функция $f(x, y) = x + y$ определена всюду, следовательно, поле направлений для данного уравнения можно построить во всей плоскости. Для того чтобы упорядочить расстановку направлений поля, воспользуемся методом изоклин.

Изоклиной поля направлений называется геометрическое место точек, в которых направление поля одинаково.

Обозначим через α угол наклона к оси OX направления поля: $\operatorname{tg} \alpha = y'$.

Изоклина, в точках которой $\alpha = 0$, т. е. $y' = \operatorname{tg} 0 = 0$, имеет уравнение $x + y = 0$.

Изоклина, в точках которой $\alpha = \frac{\pi}{6}$, т. е. $y' = \operatorname{tg} \frac{\pi}{6} = \frac{\sqrt{3}}{3}$, имеет уравнение $x + y = \frac{\sqrt{3}}{3}$ и т. д.

Чтобы по данному полю направлений начертить интегральную кривую, надо взять на плоскости XOY любую точку (x_0, y_0) за начальную точку и вести через нее линию так, чтобы она в каждой своей точке шла в направлении поля (рис. 320).

Примечание. Практически метод изоклин может быть использован для приближенного построения семейства интегральных кривых уравнения (2). При этом, чем больше построено изоклин, т. е. чем «гуще»

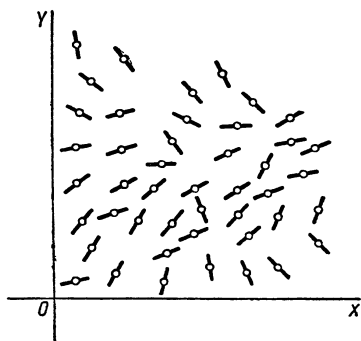


Рис. 319.

¹⁾ Если такая кривая существует.

²⁾ Мы предполагаем, что функция $f(x, y)$ в каждой точке рассматриваемой области принимает конечные значения. Тем самым исключаются из рассмотрения направления поля, параллельные оси OY .

указаны на чертеже направления поля, тем точнее можно провести интегральные кривые уравнения.

Метод изоклин позволяет представить взаимное расположение интегральных кривых уравнения.

Рассмотренные выше примеры, а также геометрическая интерпретация дифференциального уравнения первого порядка показывают, что дифференциальное уравнение первого порядка вида $y' = f(x, y)$ имеет, вообще говоря, не одно, а бесчисленное множество решений. Для того чтобы из этого бесчисленного множества решений выделить определенное решение, обычно приходится задавать значение искомой функции y_0 при некотором значении аргумента x_0 . Пару чисел x_0, y_0 называют *начальными условиями*, или *начальными данными*, решения. Геометрически зада-

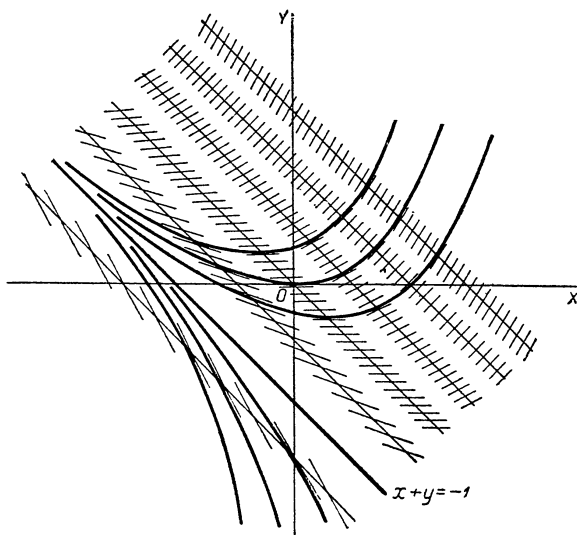


Рис. 320.

ние начальных условий равносильно заданию точки (x_0, y_0) — «начальной точки» на плоскости XOY . Будем говорить, что решение $y = \varphi(x)$ уравнения $y' = f(x, y)$ удовлетворяет начальным условиям x_0, y_0 , если $\varphi(x_0) = y_0$, т. е. если график этого решения проходит через точку (x_0, y_0) ¹⁾.

Отыскание решения дифференциального уравнения $y' = f(x, y)$, удовлетворяющего заданным начальным условиям x_0, y_0 , является одной из важнейших задач теории дифференциальных уравнений. Эта задача называется *задачей Коши*.

Естественно возникает вопрос, всегда ли существует решение задачи Коши и, если существует, будет ли оно единственным.

Ответы на эти вопросы дает теорема Коши — **теорема существования и единственности решения дифференциального уравнения первого порядка, разрешенного относительно производной**. Приведем ее формулировку.

Теорема Коши. Если функция $f(x, y)$ непрерывна в некоторой области D плоскости XOY и имеет в этой области непрерывную частную производную по y , $f'_y(x, y)$, то, какова бы ни была точка (x_0, y_0) области D , существует, и притом единственное, решение $y = \varphi(x)$ уравнения $y' = f(x, y)$, определенное в некотором интервале, содержащем точку x_0 , принимающее при $x = x_0$ значение $\varphi(x_0) = y_0$.

¹⁾ Вместо того чтобы говорить: «график решения $y = \varphi(x)$ проходит через точку», будем говорить короче: «решение $y = \varphi(x)$ проходит через точку».

Геометрически это утверждение означает, что через каждую внутреннюю точку (x_0, y_0) области D проходит, и притом только одна, интегральная кривая уравнения (рис. 321).

Из теоремы Коши следует, что в области D уравнение $y' = f(x, y)$ имеет бесчисленное множество решений. Действительно, считая x_0 постоянным и изменяя значение y_0 в некоторых пределах, получим для каждого значения y_0 свое решение: $y = \varphi(x, y_0)$ (рис. 322).

Введем теперь следующие основные определения.

Функция $y = \varphi(x, C)$, зависящая от одной произвольной постоянной C , называется *общим решением* уравнения $y' = f(x, y)$ в некоторой области σ , если она является решением этого уравнения для любого значения

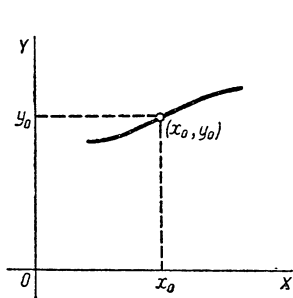


Рис. 321.

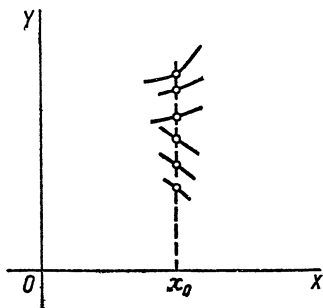


Рис. 322.

постоянной C (или для любого значения C из некоторого множества) и если любое решение уравнения, лежащее в области σ , может быть записано в виде $y = \varphi(x, C)$ при конкретном значении C .

Равенство $\Phi(x, y, C) = 0$, неявно задающее общее решение, называется *общим интегралом* уравнения (1) в области σ .

Решения, получающиеся из общего при определенных значениях постоянной C , называются *частными решениями* данного уравнения. Аналогично определяются *частные интегралы*.

Геометрически общее решение (общий интеграл) уравнения (1) в области σ представляет собой семейство интегральных кривых уравнения, обладающее тем свойством, что каждая интегральная кривая, лежащая в области σ , принадлежит этому семейству кривых¹⁾.

Если при этом функция $f(x, y)$, стоящая в правой части уравнения (2), в области σ удовлетворяет условиям теоремы Коши, то через каждую точку (x_0, y_0) области σ проходит одна определенная кривая семейства $y = \varphi(x, C)$ ($\Phi(x, y, C) = 0$). Соответствующее ей значение постоянной C можно определить, подставив координаты начальной точки (x_0, y_0) в уравнение $y = \varphi(x, C)$ ($\Phi(x, y, C) = 0$). Если $y_0 = \varphi(x_0, C_0)$ ($\Phi(x_0, y_0, C_0) = 0$), то интегральная кривая, проходящая через точку (x_0, y_0) , будет иметь уравнение $y = \varphi(x, C_0)$ ($\Phi(x, y, C_0) = 0$). Эта кривая будет изображать частное решение уравнения (1), соответствующее значению C_0 . Таким образом, для выделения из общего решения (интеграла) частного решения (интеграла) уравнения (2) достаточно задать его начальные условия.

Пример. Уравнение $y' = 2\sqrt{y}$ разрешено относительно производной. Функции $f(x, y) = 2\sqrt{y}$ и $f'_y(x, y) = \frac{1}{\sqrt{y}}$ определены и непрерывны для $y > 0$. Следовательно, условия теоремы Коши для уравнения $y' = 2\sqrt{y}$ выполнены в верхней полу-

¹⁾ Такое семейство кривых называют *однопараметрическим* с параметром C . Каждому значению C (из некоторого множества) соответствует кривая семейства, определяемая уравнением $y = \varphi(x, C)$.

плоскости $y > 0$. Через каждую ее точку проходит одно определенное решение уравнения.

Проверим, что в этой области функция $y = (x - C)^2$, где $x - C > 0$, является общим решением уравнения.

Действительно, для любого значения C

$$y' = 2(x - C), \quad 2\sqrt{y} = 2(x - C), \quad \text{т. е. } y' \equiv 2\sqrt{y}.$$

Задавая любые начальные условия $x_0, y_0, y_0 > 0$, можно подобрать значение C_0 :

$$y_0 = (x_0 - C_0)^2, \quad C_0 = x_0 - \sqrt{y_0}$$

и выделить из совокупности решений $y = (x - C)^2$ решение $y = (x - x_0 + \sqrt{y_0})^2$, удовлетворяющее заданным начальным условиям.

Геометрически общее решение данного уравнения в верхней полуплоскости представляет собой семейство правых «ветвей» парабол $y = (x - C)^2$, $x - C > 0$, т. е. $\sqrt{y} = x - C$ (рис. 323). Каждая из этих «ветвей» изображает частное решение уравнения. Задавая начальные условия $x_0, y_0, y_0 > 0$, можно из всего семейства выделить «ветвь» той параболы, которая проходит через точку (x_0, y_0) .

Примечание. Следует заметить, что найденные частные решения уравнения не исчерпывают всех его решений. Так, например, если рассматривать уравнение

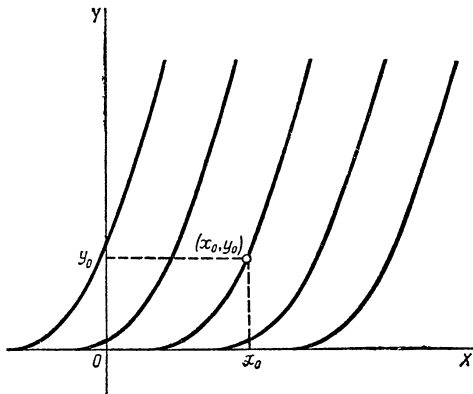


Рис. 323.

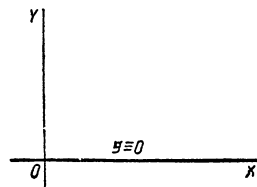


Рис. 324.

$y' = 2\sqrt{y}$ в области $y \geq 0$ (а именно в этой области определена правая часть уравнения), то к найденному выше семейству интегральных кривых $y = (x - C)^2$, $x - C > 0$, добавится решение $y \equiv 0$ (рис. 324), которое ни при каком значении постоянной C не получается из соотношения $y = (x - C)^2$ ¹⁾.

Это значит, что в области $y \geq 0$ функция $y = (x - C)^2$, где C — произвольная постоянная, уже не является общим решением данного уравнения.

Основной задачей теории дифференциальных уравнений является отыскание всех решений данного дифференциального уравнения (в явной или неявной форме) и изучение свойств этих решений. Нахождение решений дифференциальных уравнений называют *интегрированием* этих уравнений.

§ 3. Некоторые дифференциальные уравнения первого порядка, интегрируемые в квадратурах

Рассмотрим некоторые, наиболее важные типы дифференциальных уравнений первого порядка, интегрирование которых сводится к вычислению одного или нескольких неопределенных интегралов. Действие вычисления неопределенного интеграла, чтобы не было путаницы с термином «интегрирование уравнения», будем называть *квадратурой*.

¹⁾ Интегральные кривые данного уравнения в области $y \geq 0$ можно было бы также получить и путем «склеивания» частей оси OX ($y = 0$) с касающимися оси «ветвями» парабол:

$$y = \begin{cases} 0, & x < C_0, \\ (x - C_0)^2, & x \geq C_0, \end{cases} \quad C_0 \text{ — любое число.}$$

1. Уравнения вида $y' = f(x)$, где $f(x)$ — функция, определенная и непрерывная на некотором интервале $a < x < b$ оси OX ¹⁾.

Как известно из курса интегрального исчисления, все решения этого простейшего дифференциального уравнения исчерпываются соотношением

$$y = \int f(x) dx + C, \quad (1)$$

где C — произвольная постоянная²⁾.

Геометрически это значит, что все интегральные кривые уравнения $y' = f(x)$ в полосе $\{a < x < b, -\infty < y < +\infty\}$ получаются из одной из них, например $y = \int f(x) dx$, сдвигом, параллельным оси OY . Задавая в этой полосе любую точку $M_0(x_0, y_0)$, можно единственным образом определить постоянную C_0 так, что соответствующая интегральная кривая

$$y = \int f(x) dx + C_0$$

проходит через эту точку.

Соотношение (1) представляет собой общее решение уравнения $y' = f(x)$ в данной полосе.

Пример. Правая часть уравнения $y' = 3x^2$ непрерывна в промежутке $-\infty < x < +\infty$.

Общее решение уравнения во всей плоскости HOY имеет вид $y = x^3 + C$, где C — произвольная постоянная.

Найдем частное решение, удовлетворяющее начальным условиям $x_0 = 1, y_0 = 3$. Для него $3 = 1 + C_0, C_0 = 2$, т. е. $y = x^3 + 2$.

Геометрически это значит, что из семейства кубических парабол $y = x^3 + C$, представляющего собой общее решение уравнения, выделена парабола, проходящая через точку $(1, 3)$ — частное решение уравнения (рис. 325).

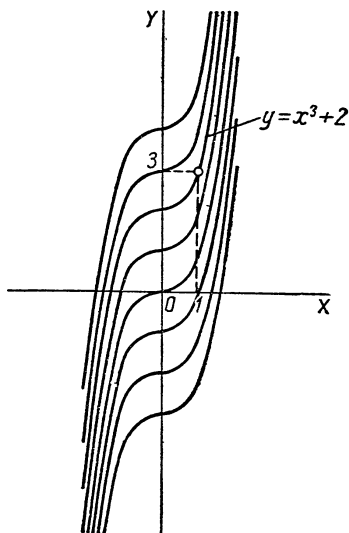


Рис. 325.

2. Уравнения с разделяющимися переменными. Дифференциальное уравнение

$y' = f(x, y)$ называется *уравнением с разделяющимися переменными*, если его можно записать в виде:

$$y' = \varphi(x) \cdot \psi(y), \quad (2)$$

т. е. если его правая часть есть произведение двух функций, одна из которых не зависит от y , а другая от x .

Предположим, что функции $\varphi(x)$ и $\psi(y)$ непрерывны на интервалах $a < x < b, c < y < d$ и что

$$\psi(y) \neq 0.$$

Умножая обе части уравнения (2) на dx и деля на $\psi(y)$, запишем его в виде:

$$\frac{dy}{\psi(y)} = \varphi(x) dx. \quad (3)$$

При этом переменные разделяются: в одной части уравнения оказывается функция от y и дифференциал y , в другой — функция от x и дифференциал x .

Пусть $\Psi(y) = \int \frac{dy}{\psi(y)}, \Phi(x) = \int \varphi(x) dx$ на рассматриваемых интервалах.

¹⁾ В частности, этот промежуток может быть бесконечным.

²⁾ Здесь, как и всюду в дальнейшем в теории дифференциальных уравнений, неопределенный интеграл обозначает одну из первообразных (а не совокупность первообразных).

Если y — решение уравнения (2), то, как следует из соотношения (3), на интервале $a < x < b$

$$d\Psi(y) \equiv d\Phi(x),$$

а значит, $\Psi(y) \equiv \Phi(x) + C$, где C — некоторое число.

Наоборот, если $\Psi(y) \equiv \Phi(x) + C$, где C — любое число, то $d\Psi(y) \equiv d\Phi(x)$, т. е. y — решение уравнения (2).

Таким образом, соотношение $\Psi(y) = \Phi(x) + C$, т. е.

$$\int \frac{dy}{\psi(y)} = \int \varphi(x) dx + C, \quad (4)$$

где C — произвольная постоянная, охватывает все решения уравнения (2) в области $\{a < x < b, c < y < d\}$. Соотношение (4) представляет собой общий интеграл уравнения (2) в указанной области.

Пример. Уравнение $y' = x(y^2 + 1)$ есть уравнение с разделяющимися переменными. Функции $\varphi(x) = x$ и $\psi(y) = y^2 + 1$ непрерывны всюду, $y^2 + 1 \neq 0$.

Разделяя переменные

$$\frac{dy}{y^2 + 1} = x dx$$

и интегрируя, получаем:

$$\operatorname{arctg} y = \frac{x^2}{2} + C \quad (5)$$

— общий интеграл данного уравнения во всей плоскости XOY .

Разрешая соотношение (5) относительно y , находим общее решение уравнения в виде

$$y = \operatorname{tg} \left(\frac{x^2}{2} + C \right), \quad -\frac{\pi}{2} < \frac{x^2}{2} + C < \frac{\pi}{2}.$$

Задавая любые начальные условия x_0, y_0 , можно из соотношения (5) определить C_0 :

$$\operatorname{arctg} y_0 = \frac{x_0^2}{2} + C_0, \quad C_0 = \operatorname{arctg} y_0 - \frac{x_0^2}{2}$$

и, следовательно, определить соответствующий частный интеграл данного уравнения:

$$\operatorname{arctg} y = \frac{x^2}{2} + \operatorname{arctg} y_0 - \frac{x_0^2}{2}$$

и частное решение:

$$y = \operatorname{tg} \left(\frac{x^2}{2} + \operatorname{arctg} y_0 - \frac{x_0^2}{2} \right).$$

Примечание. Выше предполагалось, что на интервале $c < y < d$ функция $\psi(y)$ отлична от нуля. Если для $y = y_1$, $c < y_1 < d$, имеем $\psi(y_1) = 0$, то функция — константа $y = y_1$ также является решением уравнения (2), в чем легко убедиться, подставляя $y = y_1$ в уравнение. Решение $y = y_1$ не входит в соотношение (4), т. е. соотношение (4) в этом случае уже не будет общим интегралом уравнения (2) в рассматриваемой области.

Уравнения с разделяющимися переменными иногда задают в симметричной относительно x и y дифференциальной форме:

$$M(x) N(y) dx + P(x) Q(y) dy = 0, \quad (6)$$

где функции $M(x)$, $N(y)$, $P(x)$, $Q(y)$ непрерывны соответственно в интервалах $a < x < b$, $c < y < d$.

Для нахождения всех решений такого уравнения в области $\{a < x < b, c < y < d\}$ достаточно разделить обе его части на произведение $N(y) \cdot P(x)$ и проинтегрировать полученное соотношение:

$$\int \frac{M(x)}{P(x)} dx + \int \frac{Q(y)}{N(y)} dy = C. \quad (7)$$

Если в интервалах $a < x < b$, $c < y < d$ функции $P(x)$ и $N(y)$ отличны от нуля, то соотношение (7) является общим интегралом уравнения.

Если в уравнении (6) y считается искомой функцией, а x — аргументом и $N(y_1) = 0$, $c < y_1 < d$, то $y = y_1$ — решение уравнения (6), не входящее в соотношение (7).

Если в уравнении (6) x — искомая функция, а y — аргумент и $P(x_1) = 0$, $a < x_1 < b$, то $x = x_1$ — решение уравнения (6), не входящее в соотношение (7).

3. Однородные уравнения. Функция $f(x, y)$ называется *однородной n -го измерения* относительно своих аргументов x и y , если для любого значения t имеет место тождество

$$f(tx, ty) = t^n f(x, y). \quad (8)$$

Например, $f(x, y) = x^3 + 3x^2y$ — однородная функция третьего измерения относительно x и y , так как

$$f(tx, ty) = (tx)^3 + 3(tx)^2(ty) = t^3(x^3 + 3x^2y) = t^3 f(x, y).$$

Аналогично устанавливается, что функции

$$f(x, y) = \frac{x^3 + y^3}{x^2 + xy + y^2}, \quad \varphi(x, y) = \frac{x - y}{x + 2y}, \quad \psi(x, y) = \frac{x^3}{y} + y^2 \ln \frac{x}{y}$$

являются однородными функциями соответственно первого, нулевого и второго измерений.

Функции $x^3 - 3x^2y + y^3$, $e^{x-y} + 2$, $x \sin \frac{x}{y}$ однородными не являются, так как для них условие (8) не выполняется ни при каких n .

Однородные функции, как следует из определения, обладают следующими свойствами:

1. Сумма однородных функций одинакового измерения есть однородная функция того же измерения.

2. Произведение однородных функций есть однородная функция, измерение которой равно сумме измерений сомножителей.

3. Частное однородных функций есть однородная функция. Ее измерение равно разности измерений делимого и делителя.

Дифференциальное уравнение $y' = f(x, y)$ называется *однородным*, если его правая часть $f(x, y)$ есть однородная функция нулевого измерения относительно своих аргументов.

Однородным будет также всякое уравнение вида:

$$P(x, y) dx + Q(x, y) dy = 0,$$

где $P(x, y)$ и $Q(x, y)$ — однородные функции одинакового измерения. В этом легко убедиться, разрешая уравнение относительно y' (или x').

Покажем, что интегрирование однородного уравнения

$$y' = f(x, y) \quad (9)$$

с помощью специальной подстановки сводится к интегрированию уравнения с разделяющимися переменными.

Действительно, так как $f(x, y)$ — однородная функция нулевого измерения, то для любого t

$$f(tx, ty) = f(x, y).$$

В частности, полагая $t = \frac{1}{x}$, получаем:

$$f(x, y) = f\left(1, \frac{y}{x}\right).$$

Это значит, что правая часть уравнения (9) фактически зависит от одного аргумента — отношения $\frac{y}{x}$:

$$f(x, y) = \varphi\left(\frac{y}{x}\right),$$

т. е. уравнение (9) можно записать в виде:

$$y' = \varphi\left(\frac{y}{x}\right)^{1)}.$$
 (10)

Вводя новую неизвестную функцию u с помощью подстановки $y = ux$, получим вместо (10) уравнение

$$u'x + u = \varphi(u)$$

или, что то же, уравнение

$$u' = \frac{\varphi(u) - u}{x}. \quad (11)$$

Это — уравнение с разделяющимися переменными относительно неизвестной функции u .

Предположим, что функция $\varphi(u)$ непрерывна на интервале $a < u < b$ и $\varphi(u) - u \neq 0$.

Разделяя переменные в уравнении (11) и интегрируя, находим общие интегралы этого уравнения в областях $\{a < u < b, x > 0\}$ и $\{a < u < b, x < 0\}$ в виде:

$$\int \frac{du}{\varphi(u) - u} = \int \frac{dx}{x} + C,$$

где C — произвольная постоянная.

Заменяя в этом соотношении вспомогательную функцию u ее выражением через x и y , находим в квадратурах общие интегралы данного

уравнения в областях $\{a < \frac{y}{x} < b, x > 0\}$ и $\{a < \frac{y}{x} < b, x < 0\}$ (рис. 326).

Примечание. Если $\varphi(u) \equiv u$ на интервале $a < u < b$, то уравнение (9) имеет вид $y' = \frac{y}{x}$, т. е. само является уравнением с разделяющимися переменными.

Если же $\varphi(u) - u \neq 0$ на интервале $a < u < b$, но есть значение $u = u_1$, $a < u_1 < b$, такое, что $\varphi(u_1) - u_1 = 0$, то $u = u_1$ — решение уравнения (11), а значит, $y = u_1 x$ — решение уравнения (10), не входящее в найденную выше совокупность решений.

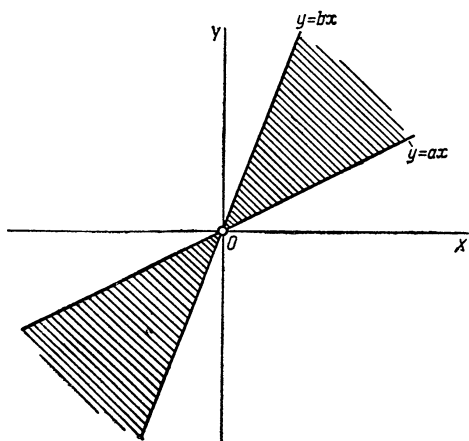


Рис. 326.

Пример. Уравнение $y' = \frac{y}{x} \left(\ln \frac{y}{x} + 1 \right)$ — однородное. Функция $f(x, y) = \frac{y}{x} \left(\ln \frac{y}{x} + 1 \right)$ определена в областях $\{x < 0, y < 0\}$ и $\{x > 0, y > 0\}$ (там $\frac{y}{x} > 0$, т. е. $\ln \frac{y}{x}$ имеет смысл).

Полагаем $\frac{y}{x} = u$, $y = ux$. При этом $y' = u'x + u$, $u'x + u = u(\ln u + 1)$,
 $u'x = u \ln u$

— уравнение с разделяющимися переменными относительно u . Решая его в областях $\{u > 0, x > 0\}$ и $\{u > 0, x < 0\}$, получаем:

$$\frac{du}{u \ln u} = \frac{dx}{x}, \quad \ln |\ln u| = \ln |x| + \ln |C|, \quad C \neq 0,$$

$$\ln u = Cx, \quad u = e^{Cx}.$$

¹⁾ При этом значения x , равные нулю, не рассматриваются.

Подставляя $u = \frac{y}{x}$, находим $\frac{y}{x} = e^{Cx}$, $y = xe^{Cx}$ — совокупность решений данного уравнения. Здесь C — любое отличное от нуля число.

При разделении переменных утеряно решение $u = 1$, т. е.

$$y = x.$$

Так как его можно получить в виде $y = xe^{Cx}$ при $C = 0$, заключаем, что $y = xe^{Cx}$, где C — любое число, — общее решение данного уравнения в областях $\{x > 0, y > 0\}$ и $\{x < 0, y < 0\}$.

Примечание. Некоторые однородные уравнения проще решаются с помощью подстановки $x = uy$, которая, как легко убедиться, также приводит однородное уравнение к уравнению с разделяющимися переменными.

4. Линейные уравнения. Дифференциальное уравнение $y' = f(x, y)$ называется *линейным*, если оно линейно относительно искомой функции y и ее производной y' , т. е. если оно может быть записано в виде:

$$y' + P(x)y = Q(x). \quad (12)$$

Примеры линейных уравнений:

$$y' + x^2y = x^3, \quad y' + x + e^x y = 0, \quad y' = y \text{ и т. д.}$$

Если в уравнении (12) правая часть $Q(x)$ не равна тождественно нулю, то это уравнение называется *линейным неоднородным* уравнением, или линейным уравнением *с правой частью*. Если же $Q(x) \equiv 0$, то уравнение (12) называется *линейным однородным* уравнением, или уравнением *без правой части*.

Уравнение $y' + P(x)y = 0$, полученное из уравнения (12) заменой функции $Q(x)$ нулем, называется линейным однородным уравнением, соответствующим данному неоднородному уравнению.

Будем рассматривать линейное уравнение

$$y' + P(x)y = Q(x)$$

на интервале $a < x < b$ непрерывности функций $P(x)$ и $Q(x)$.

Покажем, что такое уравнение интегрируется в квадратурах.

Возьмем сначала линейное однородное уравнение

$$y' + P(x)y = 0, \quad (13)$$

соответствующее данному неоднородному.

Это уравнение с разделяющимися переменными. Разделя переменные и интегрируя, находим:

$$\frac{dy}{y} = -P(x) dx, \quad \ln |y| = - \int P(x) dx + \ln |C|, \\ y = Ce^{-\int P(x) dx},$$

где C — произвольная постоянная, отличная от нуля.

При разделении переменных утеряно решение $y = 0$. Так как его можно получить из соотношения $y = Ce^{-\int P(x) dx}$ при $C = 0$, заключаем, что соотношение

$$y = Ce^{-\int P(x) dx}, \quad (14)$$

где C — произвольная постоянная, есть общее решение уравнения (13) в полосе $\{a < x < b, -\infty < y < +\infty\}$.

Для отыскания решений линейного неоднородного уравнения (12) применим *метод вариации произвольной постоянной*. А именно, будем искать решение уравнения (12) в том же виде (14), что и решение соответствующего однородного уравнения. Тогда C придется считать не постоян-

ной, а функцией от x , $C = C(x)$. Эта функция $C(x)$ должна быть такова, чтобы при подстановке

$$y = C(x) e^{-\int P(x) dx} \quad \text{и} \quad y' = C'(x) e^{-\int P(x) dx} - C(x) P(x) e^{-\int P(x) dx}$$

в уравнение (12), оно обращалось в тождество на интервале $a < x < b$.

Для определения функции $C(x)$ получаем уравнение с разделяющимися переменными:

$$C'(x) e^{-\int P(x) dx} = Q(x).$$

Интегрируя его, находим:

$$C(x) = \int Q(x) e^{\int P(x) dx} dx + C,$$

где C — произвольная постоянная.

Таким образом, для любого значения постоянной C функция

$$y = e^{-\int P(x) dx} \left(\int Q(x) e^{\int P(x) dx} dx + C \right) \quad (15)$$

является решением уравнения (12).

Наоборот, так как функция $e^{-\int P(x) dx}$ отлична от нуля, любое решение уравнения (12) на промежутке $a < x < b$ можно записать в виде:

$$y = C(x) e^{-\int P(x) dx},$$

а значит, в виде (15) при некотором значении постоянной C .

Следовательно, соотношение (15) является общим решением уравнения (12) в области $\{a < x < b, -\infty < y < +\infty\}$. Заметим, что общее решение линейного неоднородного уравнения (12)

$$y = e^{-\int P(x) dx} \left(\int Q(x) e^{\int P(x) dx} dx + C \right)$$

оказывается равным сумме общего решения соответствующего однородного уравнения ($C e^{-\int P(x) dx}$) и частного решения данного неоднородного уравнения ($e^{-\int P(x) dx} \cdot \int Q(x) e^{\int P(x) dx} dx$).

Рассмотренный выше метод вариации произвольной постоянной применяется обычно при отыскании решений линейных неоднородных уравнений.

Пример. $y' - 2xy = (x+1)e^{x^2}$ — линейное неоднородное уравнение. Функции $P(x) = -2x$ и $Q(x) = (x+1)e^{x^2}$ непрерывны всюду.

Решаем сначала линейное однородное уравнение $y' - 2xy = 0$, соответствующее данному уравнению:

$$\frac{dy}{y} = 2x dx, \quad \ln|y| = x^2 + \ln|C|, \quad y = C e^{x^2}.$$

Ищем общее решение данного уравнения в виде:

$$y = C(x) e^{x^2}.$$

Тогда $y' = C'(x) e^{x^2} + C(x) \cdot 2x e^{x^2}$. Подставляя y и y' в уравнение $y' - 2xy = (x+1)e^{x^2}$, после приведения подобных членов, получим:

$$C'(x) e^{x^2} = (x+1) e^{x^2},$$

откуда $C'(x) = x+1$, $C(x) = \frac{(x+1)^2}{2} + C$, где C — произвольная постоянная.

¹⁾ Действительно, $y = \frac{y}{e^{-\int P(x) dx}} \cdot e^{-\int P(x) dx} = C(x) e^{-\int P(x) dx}$, где $C(x) =$

$$= \frac{y}{e^{-\int P(x) dx}}.$$

²⁾ Это решение получается из общего при $C=0$.

Общее решение данного уравнения во всей плоскости XOY имеет вид:

$$y = \left[\frac{(x+1)^2}{2} + C \right] e^{x^2}.$$

5. Уравнения, приводимые к линейным. При интегрировании некоторых дифференциальных уравнений первого порядка можно использовать следующее замечание.

Если функция $y = \varphi(x, C)$ в некоторой области является общим решением уравнения

$$y' = f(x, y),$$

то выражение $v(y) = \varphi(x, C)$ является общим интегралом дифференциального уравнения

$$v'(y) y' = f(x, v(y)) \quad (16)$$

в той же области.

В справедливости этого замечания легко убедиться с помощью подстановки $z = v(y)$ в уравнение (16).

В частности, если $y' = f(x, y)$ — линейное дифференциальное уравнение:

$$y' + P(x)y = Q(x), \quad (12)$$

а $v(y)$ — любая дифференцируемая функция y , то общий интеграл уравнения

$$v'(y) y' + P(x)v(y) = Q(x) \quad (17)$$

в некоторой области запишется в виде $v(y) = \varphi(x, C)$, где $y = \varphi(x, C)$ — общее решение уравнения (12) в той же области.

Таким образом, решение любого уравнения $v'(y) y' + P(x)v(y) = Q(x)$ с помощью подстановки $v(y) = z$ приводится к решению линейного уравнения.

Так, например, уравнение Бернулли, т. е. уравнение

$$y' + P(x)y = Q(x)y^\alpha, \quad (18)$$

где α — любое действительное, отличное от нуля и от единицы число¹⁾, приводится к линейному после предварительного деления обеих частей уравнения на y^α и последующей подстановки $z = y^{1-\alpha}$.

Действительно, если $z = y^{1-\alpha}$, то $z' = (1-\alpha)y^{-\alpha}y'$. Подставляя z и z' в уравнение

$$y^{-\alpha}y' + P(x)y^{1-\alpha} = Q(x),$$

полученное из данного делением на y^α , имеем:

$$\frac{1}{1-\alpha}z' + P(x)z = Q(x),$$

или, что то же,

$$z' + (1-\alpha)P(x)z = (1-\alpha)Q(x).$$

Это — линейное уравнение относительно z . Если $P(x)$ и $Q(x)$ непрерывны, оно интегрируется в квадратурах. Находя его общее решение и подставляя $z = y^{1-\alpha}$, получим совокупность решений уравнения Бернулли.

П р и м е ч а н и е. Если $\alpha > 0$, то к найденным решениям нужно добавить решение $y = 0$, утерянное при делении на y .

П р и м е р. Уравнение Бернулли $y' + xy = x^3y^3$ ($\alpha = 3$) после подстановки $y^{-2} = z$ сводится к линейному уравнению $z' - 2xz = -2x^3$. Решая его, находим $z = x^2 + 1 + Ce^{x^2}$, где C — произвольная постоянная. Совокупность решений данного уравнения: $y^{-2} = x^2 + 1 + Ce^{x^2}$ и $y = 0$.

6. Уравнения в полных дифференциалах. Рассмотрим уравнение первого порядка, заданное в дифференциальной форме:

$$M(x, y)dx + N(x, y)dy = 0. \quad (19)$$

Если левая часть уравнения (19) представляет собой полный дифференциал некоторой функции $u(x, y)$ двух переменных, т. е. если $M(x, y)dx + N(x, y)dy = du(x, y)$, то это уравнение называется уравнением в полных дифференциалах.

¹⁾ При $\alpha = 0$ уравнение $y' + P(x)y = Q(x)$ — линейное неоднородное; при $\alpha = 1$ уравнение $y' + P(x)y = Q(x)y$ — линейное однородное.

Будем предполагать, что функции $M(x, y)$, $N(x, y)$ и их частные производные $\frac{\partial M(x, y)}{\partial y}$ и $\frac{\partial N(x, y)}{\partial x}$ непрерывны в некоторой односвязной¹⁾ области σ .

Ниже (см. гл. IX, § 4) будет доказано, что для того чтобы уравнение (19) в области σ было уравнением в полных дифференциалах, необходимо и достаточно, чтобы в области σ выполнялось условие

$$\frac{\partial M(x, y)}{\partial y} = \frac{\partial N(x, y)}{\partial x}.$$

С помощью этого условия легко установить, является ли данное уравнение уравнением в полных дифференциалах.

Так, например, уравнение $(3x^2 + 10xy) dx + (5x^2 - 1) dy = 0$ будет уравнением в полных дифференциалах во всей плоскости XOY , так как функции $M(x, y) = 3x^2 + 10xy$, $N(x, y) = 5x^2 - 1$ и их частные производные $\frac{\partial M(x, y)}{\partial y} = 10x$, $\frac{\partial N(x, y)}{\partial x} = 10x$ всюду непрерывны и $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$.

Уравнение $x^2 y dx - (5xy + 1) dy = 0$ не является уравнением в полных дифференциалах, так как (при выполнении прочих условий) нет области в плоскости XOY , в которой частные производные:

$$\frac{\partial M}{\partial y} = x^2 \quad \text{и} \quad \frac{\partial N}{\partial x} = -5y$$

были бы тождественно равны.

Если уравнение

$$M(x, y) dx + N(x, y) dy = 0 \quad (19)$$

является в области σ уравнением в полных дифференциалах, то его можно в этой области записать в виде:

$$du(x, y) = 0,$$

где $u(x, y)$ — некоторая функция.

Пусть $y = \varphi(x)$ — любое решение уравнения, лежащее в области σ . Тогда $du(x, \varphi(x)) \equiv 0$, т. е. $u(x, y) = C$, где C — некоторое число.

Наоборот, для любой функции $y = \varphi(x)$, неявно заданной уравнением $u(x, y) = C$, где C — число, имеем $du(x, y) = 0$, т. е. $y = \varphi(x)$ — решение уравнения (19).

Следовательно, соотношение $u(x, y) = C$, где C — произвольная постоянная, является общим интегралом уравнения (19) в области σ .

Пример. Уравнение $[\cos(x+y) + 2] dx + [\cos(x+y) - 5] dy = 0$ является во всей плоскости XOY уравнением в полных дифференциалах.

Так как $[\cos(x+y) + 2] dx + [\cos(x+y) - 5] dy = d[\sin(x+y) + 2x - 5y]$, то данное уравнение имеет вид:

$$d[\sin(x+y) + 2x - 5y] = 0.$$

Общим интегралом этого уравнения в плоскости XOY будет выражение

$$\sin(x+y) + 2x - 5y = C,$$

где C — произвольная постоянная.

Интегрирование уравнений в полных дифференциалах сводится фактически к отысканию функции $u(x, y)$, дифференциалом которой является левая часть уравнения (19). В приведенном примере эта функция усматривалась непосредственно. В общем случае непосредственное нахождение функции $u(x, y)$ представляет некоторые трудности, в связи с чем ниже (гл. IX, § 4) будет указан специальный метод интегрирования полного дифференциала.

¹⁾ Область σ плоскости XOY называется *односвязной*, если любые две непрерывные кривые, лежащие в этой области и имеющие общее начало и общий конец, можно непрерывно деформировать друг в друга, не выходя за пределы области.

§ 4. Уравнения первого порядка, не разрешенные относительно производной

До сих пор мы ограничивались изучением уравнений первого порядка, разрешенных относительно производной. Рассмотрим теперь уравнения первого порядка, не разрешенные относительно производной, т. е. уравнения вида:

$$F(x, y, y') = 0. \quad (1)$$

Уравнение $F(x, y, y') = 0$ неявно задает в каждой точке (x, y) некоторой области плоскости XOY одно или несколько¹⁾ действительных значений y' . Если построить в каждой такой точке отрезки с угловыми коэффициентами, равными значениям y' в этой точке, то получится так называемое *поле направлений*, определяемое уравнением $F(x, y, y') = 0$.

Проинтегрировать уравнение (1) — значит в явной или неявной форме найти все его решения. Геометрически это значит найти все кривые, касательная к каждой из которых в каждой ее точке совпадает с одним из направлений поля в этой точке.

Предположим, что в некоторой области D плоскости XOY уравнение $F(x, y, y') = 0$ неявно задает m различных действительных значений y' :

$$\begin{aligned} y' &= f_1(x, y), \quad y' = f_2(x, y), \dots, \\ y' &= f_m(x, y). \end{aligned} \quad (2)$$

В этом случае поле направлений уравнения $F(x, y, y') = 0$ в области D можно рассматривать как наложение m полей уравнений (2), разрешенных относительно производной. Все решения этих уравнений являются решениями данного уравнения в области D .

Пусть $\Phi_1(x, y, C) = 0$, $\Phi_2(x, y, C) = 0, \dots, \Phi_m(x, y, C) = 0$ — общие интегралы уравнений (2) в области D ²⁾. Совокупность этих общих интегралов называют *общим интегралом* уравнения $F(x, y, y') = 0$ в области D .

Рассмотрим несколько примеров.

Пример 1. Уравнение $y'^2 - 2xy' = 0$ не разрешено относительно производной. Разрешая его, получим $y' = 0$ и $y' = 2x$. Общие решения этих уравнений во всей плоскости XOY имеют вид: $y = C$ и $y^2 = x^2 + C$ (C — произвольная постоянная). Поэтому общий интеграл данного уравнения в плоскости XOY задается двумя соотношениями: $y = C$, $y = x^2 + C$. Определяемое данным уравнением поле направлений получается наложением полей уравнений $y' = 0$ и $y' = 2x$. Через каждую точку (x_0, y_0) плоскости XOY проходят две интегральные кривые — прямая $y = y_0$ и парабола $y = x^2 + C_0$, где $C_0 = y_0 - x_0^2$ (рис. 327).

Пример 2. Уравнение $y'^3 - 1 = 0$ не разрешено относительно производной. Разрешая его относительно y' , получаем три значения:

$$y' = 1, \quad y' = \frac{-1 + \sqrt[3]{3}i}{2}, \quad y' = \frac{-1 - \sqrt[3]{3}i}{2}$$

— одно действительное и два мнимых.

Так как нас интересуют только действительные решения уравнения, рассматриваем лишь решение $y' = 1$. Его общее решение во всей плоскости XOY имеет вид: $y = x + C$, где C — произвольная постоянная. Следовательно, все (действительные) решения данного уравнения охватываются соотношением

$$y = x + C,$$

где C — произвольная постоянная.

¹⁾ И даже, может быть, бесчисленное множество.

²⁾ Для каждого из этих интегралов произвольная постоянная C может меняться в своих пределах.

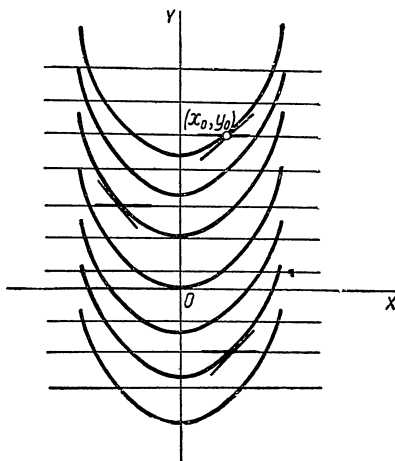


Рис. 327.

Через каждую точку (x_0, y_0) плоскости XOY проходит одна интегральная кривая данного уравнения:

$$y = x + C_0,$$

где $C_0 = y_0 - x_0$ (рис. 328).

Пример 3. Уравнение $y'^3 - 4yy' = 0$ определяет в каждой точке верхней полуплоскости ($y > 0$) три действительных значения y' :

$$y' = 0, \quad y' = 2\sqrt{y}, \quad y' = -2\sqrt{y}.$$

Решая эти уравнения, получим их общие интегралы в верхней полуплоскости:

$$y = C, \quad \sqrt{y} = x + C \quad (x + C > 0), \quad \sqrt{y} = -x + C \quad (-x + C > 0),$$

где C — произвольная постоянная.

Следовательно, в верхней полуплоскости общий интеграл данного уравнения задается соотношениями:

$$y = C, \quad \sqrt{y} = x + C \quad (x + C > 0), \quad \sqrt{y} = -x + C \quad (-x + C > 0),$$

где C — произвольная постоянная. Через каждую точку (x_0, y_0) верхней полуплоскости проходят три интегральные кривые данного уравнения:

$$y = y_0, \quad \sqrt{y} = x - x_0 + \sqrt{y_0}, \quad \sqrt{y} = -x + x_0 + \sqrt{y_0}$$

(прямая и «ветви» двух парабол) (рис. 329).

В нижней полуплоскости ($y < 0$) уравнение $y'^3 - 4yy' = 0$ имеет лишь одно действительное решение относительно y' :

$$y' = 0.$$

Общее решение уравнения $y' = 0$ в нижней полуплоскости имеет вид:

$$y = C,$$

где C — произвольная постоянная. Т. е. в нижней полуплоскости общий интеграл данного уравнения имеет вид

$$y = C,$$

где C — произвольная постоянная.

Через каждую точку (x_0, y_0) нижней полуплоскости проходит одна интегральная кривая — прямая $y = y_0$.

Примечание. Найденные выше решения уравнения $y'^3 - 4yy' = 0$ не исчерпывают всех его решений. Так, например, прямая $y = 0$, разделяющая две рассмотренные выше области, будет интегральной кривой уравнения, в чем легко убедиться подстановкой в уравнение функции $y = 0$. Интегральными кривыми будут также параболы $y = (x + C)^2$, C — произвольная постоянная, полученные путем «склеивания» в точках оси OY найденных выше интегральных кривых.

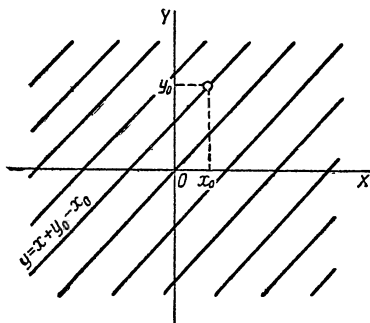


Рис. 328.

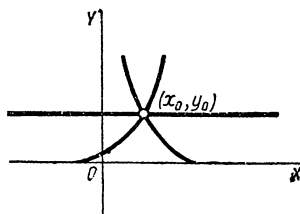


Рис. 329.

Уравнения вида $F(x, y, y') = 0$ лишь в простейших случаях удастся фактически разрешить относительно y' и тем самым свести, как это делалось выше, их решение к интегрированию уравнений, разрешенных относительно производной. В связи с этим большой интерес представляют методы, позволяющие интегрировать некоторые группы таких уравнений, не разрешая этих уравнений относительно y' .

Один из таких методов — *метод введения параметра* — будет рассмотрен ниже на примерах *неполных* дифференциальных уравнений, т. е. уравнений, не содержащих явно одной из переменных.

1. Уравнение, не содержащее явно искомой функции, имеет вид:

$$F(x, y') = 0.$$

Если это уравнение можно представить в виде:

$$x = f(y')^1), \quad (3)$$

то все решения этого уравнения могут быть найдены с помощью одной квадратуры в параметрической форме.

Действительно, будем рассматривать производную y' как параметр: $y' = t$. Пусть $y = \varphi(x)$ — любое решение уравнения (3). Выражение x через параметр t для этого решения дается самим уравнением:

$$x = f(t).$$

Найдем параметрическое выражение для y .

Так как $y' = \frac{dy}{dx}$, то $dy = y' dx$. Заменяя y' и dx их выражениями через t : $y' = t$, $dx = f'(t) dt$, получим:

$$dy = t f'(t) dt,$$

откуда, интегрируя, находим:

$$y = \int t f'(t) dt + C,$$

где C — некоторое число.

Итак, любое решение уравнения (3) может быть записано в параметрической форме в виде:

$$\begin{cases} x = f(t), \\ y = \int t f'(t) dt + C, \end{cases} \quad (4)$$

где C — некоторое число.

Если из соотношений (4) исключить t , получится (в явной или неявной форме) решение $y = \varphi(x)$.

Непосредственной подстановкой в уравнение (3) убеждаемся, что для любого значения постоянной C соотношения (4) определяют решение этого уравнения.

Таким образом, соотношение (4), где C — произвольная постоянная, охватывает все решения уравнения (3).

Примечание. При решении конкретных примеров обычно используются не формулы (4), а сам метод введения параметра.

Пример 1. Уравнение $x = y' + e^{y'}$ не может быть в элементарных функциях разрешено относительно y' , но оно уже разрешено относительно x . Найдем его решение в параметрической форме.

Принимаем $y' = t$ за параметр. Тогда $x = t + e^t$, $dx = (1 + e^t) dt$, $dy = y' dx$, т. е. $dy = t(1 + e^t) dt$,

$$y = \int t(1 + e^t) dt + C = t + te^t - e^t + C,$$

где C — произвольная постоянная.

Совокупность решений данного уравнения в параметрической форме запишется в виде:

$$\begin{cases} x = t + e^t, \\ y = t + e^t(t - 1) + C, \end{cases}$$

где C — произвольная постоянная.

Пример 2. Уравнение $y'' = x$ также можно решить методом введения параметра.

Пусть $y' = t$, тогда $x = t^2$, $dx = 2t dt$, $dy = y' dx$, т. е. $dy = 2t^2 dt$, а значит, $y = \frac{2}{3} t^3 + C$, где C — любое число.

Совокупность решений уравнения $y'' = x$ в параметрической форме записи имеет вид:

$$\begin{cases} x = t^2, \\ y = \frac{2}{3} t^3 + C, \end{cases}$$

где C — произвольная постоянная.

¹⁾ Функция $f'(t)$ предполагается интегрируемой в области изменения t .

Исключая параметр t , легко получить запись решений данного уравнения в форме:

$$y = \pm \frac{2}{3} x^{\frac{3}{2}} + C.$$

Примечание. Данное уравнение $y'^2 = x$ можно решить и не вводя параметра t , так как оно легко разрешается относительно y' :

$$y' = \pm \sqrt{x}.$$

При этом $dy = \pm \sqrt{x} dx$, $y = \pm \frac{2}{3} x^{\frac{3}{2}} + C$.

Соотношения $y = \pm \frac{2}{3} x^{\frac{3}{2}} + C$ представляют собой общий интеграл уравнения $y'^2 = x$ в полуплоскости $x > 0$.

2. Уравнение, не содержащее явно независимой переменной, имеет вид:

$$F(y, y') = 0.$$

Если это уравнение можно разрешить относительно y :

$$y = f(y')^1, \quad (5)$$

то его решения могут быть найдены с помощью квадратур в параметрической форме.

Действительно, пусть $y = \varphi(x)$ — решение уравнения (5), вдоль которого $y' \neq \text{const}$. Выбираем снова y' за параметр: $y' = t^2$). Тогда для данного решения

$$y = f(t), \quad dy = f'(t) dt, \quad dx = \frac{dy}{t} = \frac{f'(t)}{t} dt,$$

а значит, $x = \int \frac{f'(t)}{t} dt + C$, где C — некоторое число.

Итак, любое решение уравнения (5) указанного вида можно записать в параметрической форме:

$$\begin{cases} x = \int \frac{f'(t)}{t} dt + C, \\ y = f(t), \end{cases} \quad (6)$$

где C — некоторое число.

Непосредственной подстановкой в уравнение (5) убеждаемся, что для любого значения постоянной C соотношения (6) определяют решение данного уравнения.

Таким образом, соотношения (6), где C — произвольная постоянная, охватывают все решения уравнения (5), кроме, может быть, решений, вдоль которых производная y' постоянна.

Если вдоль некоторого решения $y = \varphi(x)$ производная постоянна, то это решение имеет вид $y = ax + b$, где a и b — некоторые числа.

Подставляя y в уравнение (5), получаем

$$ax + b = f(a), \quad a = 0, \quad b = f(0).$$

Таким образом, к соотношениям (6) в случае, когда $f(0)$ имеет смысл, надо добавить решение $y = f(0)$.

Пример. $y = y' + \ln y'$ — неполное уравнение, не разрешенное относительно производной. Будем решать его в параметрической форме.

¹⁾ Функция $f(t)$, как следует из приведенных ниже рассуждений, должна быть такой, чтобы функция $\frac{f'(t)}{t}$ была интегрируема в области изменения параметра t .

²⁾ Если вдоль некоторого решения производная y' постоянна, ее нельзя принимать за параметр t .

Пусть $y' = t$ — параметр. Тогда

$$y = t + \ln t, \quad dy = \left(1 + \frac{1}{t}\right) dt,$$

$$dx = \frac{dy}{t}, \quad \text{т. е.} \quad dx = \left(\frac{1}{t} + \frac{1}{t^2}\right) dt,$$

а значит,

$$x = \ln t - \frac{1}{t} + C,$$

где C — произвольная постоянная.

Соотношения

$$\begin{cases} x = \ln t - \frac{1}{t} + C, \\ y = t + \ln t, \end{cases}$$

где C — произвольная постоянная, охватывают все решения данного уравнения.

Метод введения параметра можно также применять к интегрированию некоторых групп полных уравнений первого порядка, не разрешенных относительно производных.

§ 5. Особые решения

1. Понятие об особых точках и особых решениях для уравнений, разрешенных относительно производной. Пусть дано уравнение $y' = f(x, y)$. Обозначим через D область определения функции $f(x, y)$. Будем называть внутренней точку (x_0, y_0) области D *обыкновенной точкой* уравнения $y' = f(x, y)$, если существует окрестность этой точки, в которой выполнены условия теоремы Коши. Через каждую обыкновенную точку области D (в указанной окрестности) проходит, и притом одна, интегральная кривая уравнения. Задача Коши, поставленная для обыкновенной точки уравнения, имеет, и притом единственное, решение.

Граничные точки области D , а также те внутренние точки, которые не являются обыкновенными, будем называть *особыми точками* данного уравнения. Так как в любой сколько угодно малой окрестности особой точки не выполнены условия теоремы Коши, через эту точку может не проходить ни одной интегральной кривой уравнения, может проходить как одна, так и несколько и даже бесчисленное множество интегральных кривых. Задача Коши, поставленная для особой точки уравнения, может не иметь решения, может иметь как единственное, так и не единственное решение.

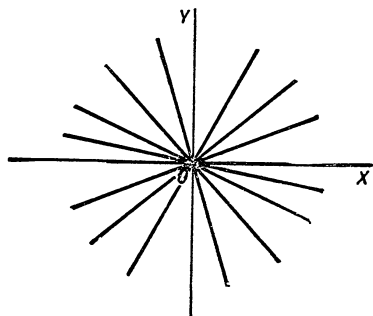


Рис. 330.

Пример 1. Для уравнения $y' = \frac{y}{x}$ функция $f(x, y) = \frac{y}{x}$ определена всюду, кроме прямой $x = 0$. В левой ($x < 0$) и правой ($x > 0$) полуплоскости выполнены условия теоремы Коши [функции $f(x, y) = \frac{y}{x}$ и $f'_y(x, y) = \frac{1}{x}$ непрерывны].

Следовательно, каждая точка (x, y) , где $x \neq 0$, — обыкновенная точка данного уравнения. Точки оси OY ($x = 0$) являются особыми точками уравнения.

Совокупность решений уравнения $y' = \frac{y}{x}$ с разделяющимися переменными: $y = Cx$, где C — произвольная постоянная, $x < 0$ и $x > 0$, геометрически представляет собой семейство лучей, выходящих из начала координат. При этом из особой точки $(0, 0)$ выходит бесчисленное множество интегральных кривых, из остальных особых точек не выходит ни одной интегральной кривой данного уравнения (рис. 330).

Пример 2. Для уравнения $y' = 3\sqrt[3]{y^2}$ функция $f(x, y) = 3\sqrt[3]{y^2}$ определена и непрерывна во всей плоскости XOY . Ее частная производная по y , $f'_y = \frac{2}{\sqrt[3]{y}}$,

непрерывна всюду, кроме точек прямой $y=0$. Эти точки являются особыми точками данного уравнения.

Непосредственной подстановкой в уравнение $y' = 3\sqrt[3]{y^2}$ убеждаемся, что функции $y=0$ и $y=(x+C)^3$, где C — произвольная постоянная, являются решениями уравнения.

В каждой особой точке уравнения $(x_0, 0)$ нарушена единственность решения, так как через нее проходят по крайней мере две интегральные кривые — прямая $y=0$ и кубическая парабола $y=(x+C_0)^3$, где $C_0=-x_0$ (рис. 331).

В отличие от предыдущего примера геометрическое место особых точек данного уравнения представляет собой интегральную кривую уравнения.

Будем называть решение уравнения $y'=f(x, y)$ *особым*, если в каждой его точке нарушена единственность, т. е. если через каждую его точку, кроме него самого, проходит еще хотя бы одна интегральная кривая уравнения.

Для отыскания особых решений уравнения $y'=f(x, y)$ в области D достаточно найти геометрическое место особых точек уравнения, непосредственной подстановкой в уравнение проверить, является ли оно или его части интегральными кривыми уравнения, и в случае положительного ответа проверить, нарушается ли в точках этих интегральных кривых единственность.

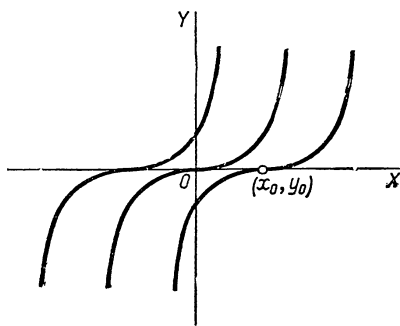


Рис. 331.

Пример 3. Для уравнения $y' = 3\sqrt[3]{y^2}$, как было установлено выше, геометрическое место особых точек — прямая $y=0$ — является особым решением уравнения.

Пример 4. Уравнение $y' = 3\sqrt[3]{y^2} + 1$ особых решений не имеет, так как геометрическое место его особых точек — та же прямая $y=0$ — интегральной кривой уравнения не является.

Пример 5. Уравнение

$$y' = \begin{cases} y \ln y, & y > 0, \\ 0, & y = 0 \end{cases}$$

особых решений не имеет, так как геометрическое место его особых точек — прямая $y=0$, хотя и является решением уравнения, но не будет особым решением (ни в одной его точке не нарушается единственности, так как остальные интегральные кривые уравнения, т. е. кривые

$$y = e^{Ce^x},$$

где C — произвольная постоянная, не пересекают прямой $y=0$).

2. Огибающая однопараметрического семейства кривых. Рассмотрим уравнение

$$\Phi(x, y, C) = 0, \quad (1)$$

где C — переменная (параметр), изменяющаяся в некоторых пределах (в частности, от $-\infty$ до $+\infty$).

Каждому значению параметра C , $C=C_0$, соответствует уравнение $\Phi(x, y, C_0)=0$, связывающее переменные x и y . Это уравнение, вообще говоря, определяет на плоскости XOY некоторую кривую. Множество всех таких кривых называется *однопараметрическим семейством кривых*, заданных уравнением (1).

Уравнение (1) называется уравнением этого однопараметрического семейства кривых.

Например, уравнение $y-x+C=0$, $-\infty < C < +\infty$, определяет семейство всех прямых плоскости XOY с угловым коэффициентом, равным единице; уравнение $x^2+y^2=C$, $0 < C < +\infty$, определяет семейство концентрических окружностей плоскости с центром $(0, 0)$ (рис. 332) и т. д.

При изучении некоторых семейств кривых существенную роль играет так называемая огибающая кривая.

Огибающей семейства кривых

$$\Phi(x, y, C) = 0$$

называется кривая, которая в каждой своей точке касается некоторой отличной от нее кривой семейства и каждого куска которой касается бесчисленное множество различных кривых семейства.

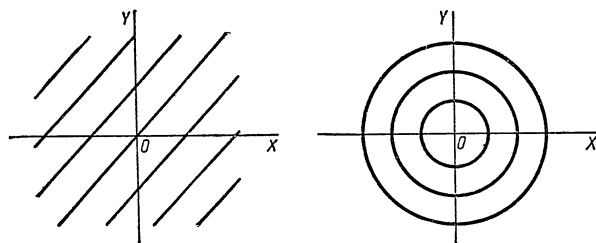


Рис. 332.

Например, семейство синусоид $y = \sin(x + C)$, $-\infty < C < +\infty$, имеет огибающую, состоящую из двух прямых $y = \pm 1$ (рис. 333), для семейства парабол $y = (x - C)^2$, $-\infty < C < +\infty$, огибающей будет ось OX (рис. 334), а для семейства окружностей $(x - C)^2 + (y - \sqrt{4 - C^2})^2 = 1$, $-2 \leq C \leq 2$, огибающая состоит из двух полуокружностей: $y = \sqrt{1 - x^2}$ и $y = -\sqrt{1 - x^2}$ (рис. 335).

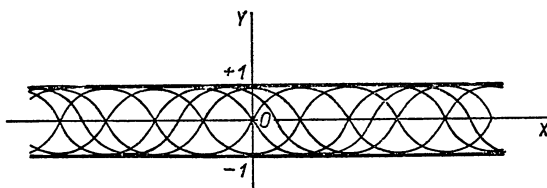


Рис. 333.

Следует заметить, что не всякое однопараметрическое семейство кривых имеет огибающую. Так, например, рассмотренные выше семейства прямых ($y - x + C = 0$) и окружностей ($x^2 + y^2 = C$) огибающих не имеют.

Предположим, что функция $\Phi(x, y, C)$, входящая в левую часть уравнения (1), имеет непрерывные частные производные по всем аргу-

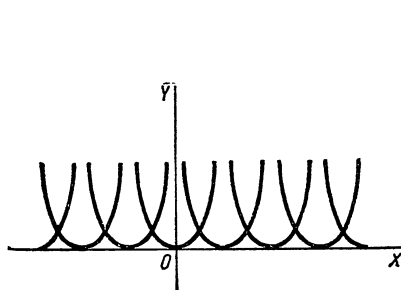


Рис. 334.

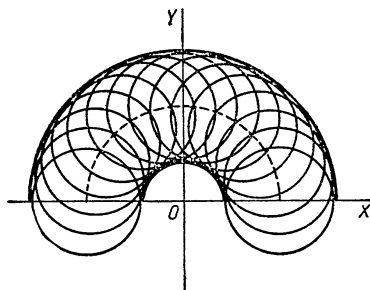


Рис. 335.

ментам. Можно показать, что если при этом семейство (1) имеет огибающую, то она входит в состав *дискриминантной кривой* семейства, т. е. кривой, уравнение которой получается после исключения параметра C из двух уравнений:

$$\begin{cases} \Phi(x, y, C) = 0, \\ \Phi'_C(x, y, C) = 0. \end{cases}$$

Помимо огибающей в состав дискриминантной кривой могут входить геометрические места *особых точек* кривых семейства, т. е. точек, в которых $\Phi'_x(x, y, C) = 0$, $\Phi'_y(x, y, C) = 0$.

Таким образом, для определения огибающей семейства кривых $\Phi(x, y, C) = 0$ достаточно продифференцировать уравнение семейства по C , и затем из уравнения семейства кривых $\Phi(x, y, C) = 0$ и полученного уравнения $\Phi'_C(x, y, C) = 0$ исключить параметр C .

Если во всех точках найденной при этом дискриминантной кривой или некоторых из составляющих ее кривых производная $\Phi'_x(x, y, C)$ или $\Phi'_y(x, y, C)$ отлична от нуля, то эти кривые будут составлять огибающую.

Если же в точках дискриминантной кривой (или кривой, входящей в ее состав) выполняются условия:

$$\Phi'_x(x, y, C) = \Phi'_y(x, y, C) = 0,$$

то соответствующая кривая есть геометрическое место особых точек и огибающей, вообще говоря, не является.

Пример 1. Дано семейство полукубических парабол:

$$(y - C)^2 = (x - C)^3, \quad -\infty < C < +\infty.$$

Запишем его уравнение в виде:

$$(y - C)^2 - (x - C)^3 = 0$$

$$(\Phi(x, y, C) = (y - C)^2 - (x - C)^3),$$

продифференцируем по C :

$$-2(y - C) + 3(x - C)^2 = 0$$

и исключим параметр C из этих соотношений:

$$\begin{aligned} (y - C)^2 - (x - C)^3 &= 0, \\ -2(y - C) + 3(x - C)^2 &= 0. \end{aligned}$$

Получим:

- 1) $x - C = 0$, $y - C = 0$, т. е. $y = x$, и
- 2) $x - C = \frac{4}{9}$, $y - C = \frac{8}{27}$, т. е. $y = x - \frac{4}{27}$.

Дискриминантная кривая состоит из двух прямых: $y = x$ и $y = x - \frac{4}{27}$.

В точках прямой $y = x$ имеем $\Phi'_x(x, y, C) = -3(x - C)^2 = 0$, $\Phi'_y(x, y, C) = 2(y - C) = 0$, т. е. эта прямая является геометрическим местом особых точек кривых семейства — вершин полукубических парабол.

В точках прямой $y = x - \frac{4}{27}$ имеем $\Phi'_x(x, y, C) = -3(x - C)^2 = -\frac{16}{27} \neq 0$. Эта прямая является огибающей данного семейства кривых (см. рис. 336).

Пример 2. Дано семейство кривых $y = C(x - C)^2$, $-\infty < C < +\infty$. Здесь $\Phi(x, y, C) = y - C(x - C)^2$.

Продифференцируем уравнение семейства по параметру C :

$$-(x - C)^2 + 2C(x - C) = 0,$$

исключим C из соотношений:

$$\begin{aligned} y &= C(x - C)^2, \\ (x - C)^2 - 2C(x - C) &= 0. \end{aligned}$$

Получим:

- 1) $x - C = 0$, $y = 0$ и
- 2) $-x + 3C = 0$, $C = \frac{1}{3}x$, $y = \frac{4}{27}x^3$, т. е. дискриминантная кривая со-

стоит из прямой $y = 0$ (ось OX) и кубической параболы $y = \frac{4}{27}x^3$.

Так как в точках прямой $y=0$ $\Phi'_y(x, y, C)=1 \neq 0$ и в точках кубической параболы $y=\frac{4}{27}x^3$ также $\Phi'_y(x, y, C)=1 \neq 0$, то эти линии составляют огибающую данного семейства кривых (рис. 337).

В данном примере прямая $y=0$, являясь составной частью огибающей, сама принадлежит рассматриваемому семейству кривых ($C=0$).

3. Особые решения дифференциальных уравнений как огибающие семейства интегральных кривых. Если однопараметрическое семейство интегральных кривых

$$\Phi(x, y, C) = 0$$

дифференциального уравнения $y' = f(x, y)$ имеет огибающую, то эта огибающая является особым решением данного уравнения.

Действительно, в каждой своей точке огибающая касается некоторой интегральной кривой. Это значит, что в каждой точке огибающей значения x , y и y' , соответствующие этой точке, удовлетворяют дифференциальному уравнению $y' = f(x, y)$, т. е. огибающая семейства интегральных кривых сама является интегральной кривой уравнения $y' = f(x, y)$.

Так как через каждую точку огибающей, кроме нее самой, касаясь, проходит еще какая-то интегральная кривая семейства, то в каждой точке огибающей нарушена единственность решения, т. е. огибающая семейства интегральных кривых изображает именно особое решение данного дифференциального уравнения.

В случае, если функция $\Phi(x, y, C)$ имеет непрерывные частные производные по всем своим аргументам, для отыскания особых решений уравнения $y' = f(x, y)$ можно использовать указанный выше прием для определения огибающей семейства кривых как составной части дискриминантной кривой семейства.

Пример 1. Для уравнения $y' = 3\sqrt[3]{y^2}$ однопараметрическое семейство интегральных кривых может быть задано уравнением

$$y = (x + C)^3, \quad -\infty < C < +\infty.$$

Дифференцируя уравнение этого семейства по C :

$$3(x + C)^2 = 0$$

и исключая C из соотношений:

$$\begin{aligned} y - (x + C)^3 &= 0, \\ 3(x + C)^2 &= 0, \end{aligned}$$

получаем дискриминантную кривую: $y=0$.

Так как в точках этой кривой $\Phi'_y(x, y, C)=1 \neq 0$, заключаем, что $y=0$ — огибающая семейства интегральных кривых, т. е. особое решение уравнения $y' = 3\sqrt[3]{y^2}$ (рис. 331).

Определение особых решений уравнений, не разрешенных относительно производной, в данном курсе не рассматривается.

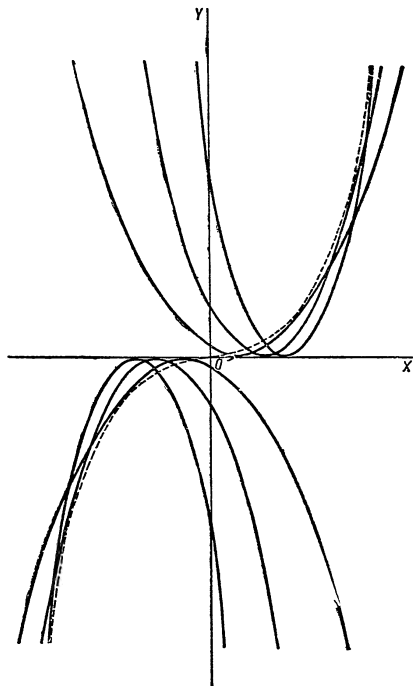


Рис. 337.

§ 6. Составление дифференциального уравнения однопараметрического семейства кривых

Поставим задачу в некотором смысле обратную задаче интегрирования дифференциального уравнения первого порядка. Дано однопараметрическое семейство кривых

$$\Phi(x, y, C) = 0,$$

требуется составить дифференциальное уравнение, для которого кривые этого семейства будут интегральными кривыми.

Координаты (x, y) точек любой кривой данного семейства удовлетворяют уравнению

$$\Phi(x, y, C) = 0 \quad (1)$$

при соответствующем этой кривой постоянном значении C .

Рассматривая это уравнение как неявное задание y как функции x для каждой кривой семейства, продифференцируем его по x :

$$\Phi'_x(x, y, C) + \Phi'_y(x, y, C)y' = 0^1. \quad (2)$$

Соотношению (2) удовлетворяют координаты точек (x, y) и угловые коэффициенты касательных в этих точках к каждой кривой семейства при соответствующем этой кривой значении C .

Если теперь из соотношений (1) и (2) исключить параметр C , то получится соотношение

$$F(x, y, y') = 0,$$

связывающее значения x , y и y' для точек любой из кривых семейства (1).

Это соотношение представляет собой дифференциальное уравнение первого порядка, вообще говоря, не разрешенное относительно производной. Оно называется *дифференциальным уравнением* данного семейства кривых. Каждая из кривых семейства (1) будет интегральной кривой этого уравнения.

Дифференциальное уравнение семейства кривых обычно используется для изучения свойств, общих всем кривым данного семейства.

Пример 1. Составим дифференциальное уравнение семейства синусоид $y = \sin(x + C)$, $-\infty < C < +\infty$ (рис. 333).

Дифференцируя уравнение семейства $y = \sin(x + C)$ по x (при постоянном значении C): $y' = \cos(x + C)$ и исключая из соотношений

$$\begin{aligned} y &= \sin(x + C), \\ y' &= \cos(x + C) \end{aligned}$$

параметр C , получим искомое дифференциальное уравнение

$$y^2 + y'^2 = 1.$$

Пример 2. Составим дифференциальное уравнение семейства

$$y - C = Ce^x, \quad -\infty < C < +\infty.$$

Дифференцируя это уравнение по x :

$$y' = Ce^x$$

и исключая C из соотношений:

$$y - C = Ce^x, \quad y' = Ce^x,$$

находим

$$y - C = y', \quad C = y - y', \quad y' = (y - y')e^x$$

— дифференциальное уравнение данного семейства кривых.

¹⁾ Функция $\Phi(x, y, C)$ предполагается удовлетворяющей всем необходимым для этих операций условиям.

§ 7. Задача об изогональных траекториях

Рассмотрим однопараметрическое семейство кривых, обладающее тем свойством, что через каждую точку некоторой области плоскости XOY проходит, и притом одна, кривая семейства.

Изогональной траекторией такого семейства кривых называется кривая, которая в каждой своей точке пересекает проходящую через эту точку кривую семейства под одним и тем же углом ω . В частности, если $\omega = \frac{\pi}{2}$, траектория называется *ортогональной*. Изогональных траекторий к данному семейству кривых существует, вообще говоря, бесчисленное множество.

Поставим своей целью по данному уравнению семейства кривых

$$\Phi(x, y, C) = 0 \quad (1)$$

и заданному углу ω найти соответствующее семейство изогональных траекторий.

Пусть $M(x, y)$ — любая точка изогональной траектории. Через эту точку проходит некоторая кривая семейства (1). Обозначим через α и φ углы наклона к оси OX касательных к изогональной траектории и кривой семейства в точке M (рис. 338). Так как по условию $\alpha = \varphi + \omega$, то $\operatorname{tg} \alpha = \operatorname{tg}(\varphi + \omega)$, т. е.

$$\operatorname{tg} \alpha = \frac{\operatorname{tg} \varphi + \operatorname{tg} \omega}{1 - \operatorname{tg} \varphi \cdot \operatorname{tg} \omega},$$

в случае, если $\omega \neq \frac{\pi}{2}$, и

$$\operatorname{tg} \alpha = -\operatorname{ctg} \varphi = -\frac{1}{\operatorname{tg} \varphi}$$

в случае, если $\omega = \frac{\pi}{2}$.

Учитывая геометрический смысл производной, заключаем, что в точке M

$$y'_{\text{из. тр.}} = \operatorname{tg} \alpha, \quad y'_{\text{кр.}} = \operatorname{tg} \varphi,$$

т. е.

$$y'_{\text{из. тр.}} = \frac{y'_{\text{кр.}} + k}{1 - ky'_{\text{кр.}}}, \quad k = \operatorname{tg} \omega, \quad \text{если } \omega \neq \frac{\pi}{2},$$

и

$$y'_{\text{из. тр.}} = -\frac{1}{y'_{\text{кр.}}}, \quad \text{если } \omega = \frac{\pi}{2}.$$

Пусть теперь $y' = f(x, y)$ — дифференциальное уравнение данного семейства кривых. Так как точка M принадлежит одной из кривых семейства (1), то в этой точке $y'_{\text{кр.}} = f(x, y)$, а значит, в той же точке

$$y'_{\text{из. тр.}} = \frac{f(x, y) + k}{1 - kf(x, y)}$$

в случае, если $\omega \neq \frac{\pi}{2}$,

и

$$y'_{\text{из. тр.}} = -\frac{1}{f(x, y)}$$

в случае, если $\omega = \frac{\pi}{2}$.

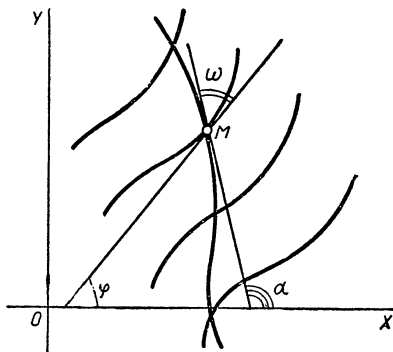


Рис. 338.

Полученные соотношения, связывающие x , y и y' для любой точки M изогональной ($\omega \neq \frac{\pi}{2}$) и ортогональной траекторий, являются дифференциальными уравнениями этих траекторий. Решая их, найдем семейства изогональных и ортогональных траекторий данного семейства кривых.

Примечание. Если дифференциальное уравнение семейства (1) найдено в виде

$$F(x, y, y') = 0,$$

не разрешенном относительно производной, то так как

$$y'_{кр.} = \frac{y'_{из.тр.} - k}{1 + ky'_{из.тр.}}$$

при $\omega \neq \frac{\pi}{2}$ и $y'_{кр.} = -\frac{1}{y'_{из.тр.}}$ при $\omega = \frac{\pi}{2}$, дифференциальные уравнения изогональных ($\omega \neq \frac{\pi}{2}$) и ортогональных траекторий соответственно имеют вид:

$$F\left(x, y, \frac{y' - k}{1 + ky'}\right) = 0,$$

$$F\left(x, y, -\frac{1}{y'}\right) = 0.$$

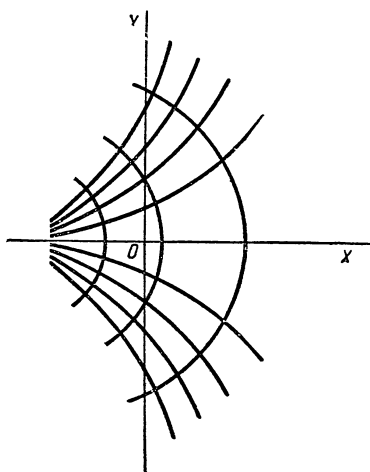


Рис. 339.

Пример 1. Дано семейство кривых $y = Ce^x$, $-\infty < C < +\infty$. Найдем их ортогональные траектории.

Для этого составляем дифференциальное уравнение данного семейства линий:

$$y = Ce^x, \quad y' = Ce^x, \quad y' = y,$$

по нему составляем дифференциальное уравнение ортогональных траекторий:

$$y' = -\frac{1}{y}.$$

Решая полученное уравнение¹⁾, находим семейство ортогональных траекторий — семейство парабол:

$$\frac{y^2}{2} = -x + C, \quad -\infty < C < +\infty \quad (\text{рис. 339}).$$

Пример 2. Дано семейство прямых, проходящих через начало координат: $y = Cx$, $-\infty < C < +\infty$. Найдем кривые, пересекающие прямые $y = Cx$ под углом $\omega = \frac{\pi}{4}$.

Для нахождения изогональных траекторий, соответствующих углу $\omega = \frac{\pi}{4}$ ($k = 1$), составляем дифференциальное уравнение семейства прямых:

$$y = Cx, \quad y' = C, \quad \text{т. е.} \quad y' = \frac{y}{x} \quad (x \neq 0),$$

по нему составляем дифференциальное уравнение изогональной траектории:

$$y' = \frac{1 + \frac{y}{x}}{1 - \frac{y}{x}}$$

¹⁾ При решении уравнения не рассматриваются значения $y = 0$.

и решаем это однородное уравнение:

$$y = ux, \quad y' = u'x + u, \quad u'x + u = \frac{1+u}{1-u}, \quad \frac{1-u}{1+u^2} du = \frac{dx}{x},$$

$$\operatorname{arctg} u - \frac{1}{2} \ln(u^2 + 1) = \ln|x| + C,$$

$$\operatorname{arctg} \frac{y}{x} - \frac{1}{2} \ln(x^2 + y^2) + \frac{1}{2} \ln x^2 = \ln|x| + C,$$

$$\operatorname{arctg} \frac{y}{x} - C = \ln \sqrt{x^2 + y^2}.$$

Переходя к полярным координатам: $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, запишем решение в виде:

$$\varphi - C = \ln \rho, \quad \rho = e^{\varphi - C}$$

или, полагая $e^{-C} = C_1$, в виде $\rho = C_1 e^{\varphi}$. Это семейство логарифмических спиралей (рис. 340).

§ 8. Приближенные методы интегрирования дифференциальных уравнений первого порядка

Известные методы точного интегрирования дифференциальных уравнений первого порядка (некоторые из них были изложены выше) пригодны лишь для сравнительно небольшой части уравнений, встречающихся на практике. В связи с этим большое значение приобретают методы приближенного интегрирования дифференциальных уравнений.

В зависимости от формы, в которой они представляют решения, эти методы делятся в основном на три группы:

1. *аналитические методы*, дающие приближенные решения дифференциальных уравнений в виде аналитических выражений;
2. *графические методы*, дающие приближенные решения уравнений в виде графиков;
3. *численные методы*, дающие приближенные численные значения решений для указанных значений аргумента.

Эта классификация является в известной степени условной. Так, например, приведенные ниже графические методы дают одновременно способы численного решения и т. д.

Рассмотрим некоторые из наиболее употребительных методов приближенного решения задачи Коши для дифференциальных уравнений первого порядка.

1. **Метод последовательных приближений.** Пусть дано уравнение

$$y' = f(x, y), \quad (1)$$

правая часть которого в прямоугольнике $\{|x - x_0| \leq a, |y - y_0| \leq b\}$ непрерывна и имеет непрерывную частную производную по y . Требуется найти решение уравнения (1), удовлетворяющее начальным условиям x_0, y_0 . По теореме Коши в некоторой окрестности точки (x_0, y_0) такое решение существует и является единственным. Рассмотрим вспомогательное уравнение

$$y = y_0 + \int_{x_0}^x f(x, y) dx \quad (2)$$

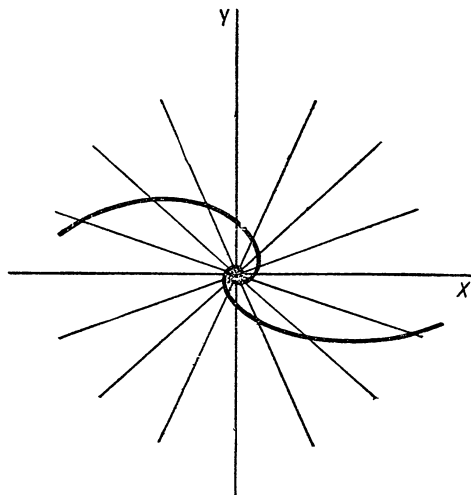


Рис. 340.

с неизвестной функцией $y=y(x)$. Такое уравнение называется *интегральным* (неизвестная функция входит под знак интеграла). Всякое непрерывное решение уравнения (2) (т. е. всякая непрерывная функция, обращающая это уравнение в тождество) является вместе с тем решением уравнения (1) с начальными условиями x_0, y_0 .

Действительно, если $y=y(x)$ — непрерывная функция, для которой

$$y(x) \equiv y_0 + \int_{x_0}^x f(x, y(x)) dx,$$

то она дифференцируема:

$$y'(x) \equiv \frac{d}{dx} \left(y_0 + \int_{x_0}^x f(x, y(x)) dx \right) \equiv f(x, y(x)) ^1)$$

и

$$y(x_0) = y_0 + \int_{x_0}^{x_0} f(x, y(x)) dx = y_0,$$

т. е. функция $y=y(x)$ удовлетворяет уравнению (1) и начальным условиям x_0, y_0 .

Таким образом, для решения поставленной выше задачи Коши достаточно найти непрерывное решение интегрального уравнения (2). С этой целью применим так называемый *метод последовательных приближений*. А именно, построим бесконечную последовательность непрерывных функций:

$$\begin{aligned} y_0(x) &= y_0, \\ y_1(x) &= y_0 + \int_{x_0}^x f(x, y_0) dx, \\ y_2(x) &= y_0 + \int_{x_0}^x f(x, y_1(x)) dx, \\ &\dots \dots \dots \\ y_n(x) &= y_0 + \int_{x_0}^x f(x, y_{n-1}(x)) dx, \\ &\dots \dots \dots \end{aligned}$$

называемых соответственно нулевым, первым, вторым, ..., n -м *приближениями* решения.

Можно показать, что для каждого значения x в некоторой окрестности точки x_0 последовательность $y_1(x), y_2(x), \dots, y_n(x), \dots$ имеет предел:

$$\lim_{n \rightarrow \infty} y_n(x) = y(x),$$

причем функция $y(x)$ непрерывна и является решением уравнения (2) в этой окрестности.

Отсюда следует, что решение $y=y(x)$ уравнения $y'=f(x, y)$ с начальными условиями x_0, y_0 в некоторой окрестности точки x_0 может быть найдено с любой степенью точности, если положить

$$y(x) \approx y_n(x),$$

где n — достаточно большое число.

Метод последовательных приближений позволяет найти аналитическое выражение приближенного решения поставленной задачи Коши, т. е. является аналитическим методом.

¹⁾ Применяется теорема о дифференцировании определенного интеграла по верхнему пределу.

Оценка погрешности, получающейся от замены решения $y(x)$ его n -м приближением в окрестности $x_0 - h < x < x_0 + h$, дается следующим неравенством:

$$|y(x) - y_n(x)| \leq \frac{MN^n h^{n+1} e^{Nh}}{(n+1)!},$$

где $M = \max |f(x, y)|$, $N = \max |f'_y(x, y)|$ в прямоугольнике $\{|x - x_0| \leq a, |y - y_0| \leq b\}$, а h — меньшее из чисел a и $\frac{b}{M}$.

Пример. Найдём методом последовательных приближений n -ое приближение решения уравнения $y' = y$ с начальными условиями $0, 1$ и сравним его с точным решением этой задачи Коши.

Функции $f(x, y) = y$ и $f'_y(x, y) = 1$ непрерывны в любом прямоугольнике $\{|x| \leq a, |y - 1| \leq b\}$.

По определению

$$\begin{aligned} y_0 &= 1, \\ y_1(x) &= 1 + \int_0^x 1 dx = 1 + x, \\ y_2(x) &= 1 + \int_0^x (1 + x) dx = 1 + x + \frac{x^2}{2!}, \\ y_3(x) &= 1 + \int_0^x \left(1 + x + \frac{x^2}{2!}\right) dx = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}, \\ &\dots\dots\dots \\ y_n(x) &= 1 + \int_0^x \left(1 + x + \dots + \frac{x^{n-1}}{(n-1)!}\right) dx = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}, \\ &\dots\dots\dots \end{aligned}$$

Пусть $a = 1$, $b = 1$. Тогда $M = 2$, $N = 1$, $h = \frac{1}{2}$.

Если положить

$$y(x) \approx 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}$$

на интервале $-\frac{1}{2} < x < \frac{1}{2}$, то допущенная при этом погрешность удовлетворяет неравенству

$$\left| y(x) - \left(1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}\right) \right| \leq \frac{2 \cdot \frac{1}{2^{n+1}} \cdot e^{1/2}}{(n+1)!}.$$

В частности, если взять $n = 4$, то погрешность равенства

$$y(x) \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}$$

на интервале $-\frac{1}{2} < x < \frac{1}{2}$ не будет превышать 0,001.

Интегрируя уравнение $y' = y$ методом разделения переменных и определяя по начальным условиям $0; 1$ постоянную C , получим точное решение поставленной задачи Коши: $y = e^x$ ¹⁾.

Ниже, в гл. X, будет рассмотрен ещё один аналитический метод приближенного решения задачи Коши, часто применяемый на практике — разложение искомого решения в степенной ряд с последующей заменой ряда его частной суммой.

¹⁾ Рекомендуется сравнить найденное приближенное решение с разложением точного решения e^x по формуле Маклорена.

2. Метод Эйлера. Как уже говорилось выше, уравнение $y' = f(x, y)$ в области определения функции $f(x, y)$ (обозначим ее через D) задает поле направлений. Найти графически решение этого уравнения с начальными данными x_0, y_0 — значит через точку $M_0(x_0, y_0)$ провести кривую, касающуюся в каждой своей точке направления поля в этой точке. Метод Эйлера позволяет приближенно построить такую кривую, не строя предварительно поля направлений во всей области D .

Пусть требуется найти (графически) решение уравнения

$$y' = f(x, y) \quad (1)$$

с начальными данными x_0, y_0 для значений x , больших x_0 (построение решения на участке, лежащем слева от x_0 , производится аналогично).

Возьмем последовательность точек $x_0, x_1, x_2, \dots$ на оси OX ($x_0 < x_1 < x_2 < \dots$) и проведем через них вертикальные прямые (рис. 341). Подставив в правую часть уравнения (1) координаты точки $M_0(x_0, y_0)$, вычислим угловой коэффициент касательной к искомой интегральной кривой в этой точке:

$$\operatorname{tg} \alpha_0 = y'|_{M_0} = f(x_0, y_0).$$

Заменяем на участке $[x_0, x_1]$ интегральную кривую отрезком касательной к ней в точке M_0 , т. е. проведем из точки M_0 прямую в направлении поля в точке M_0 до пересечения с вертикалью $x = x_1$ в некоторой точке $M_1(x_1, y_1)$. При этом

$$y_1 = y_0 + f(x_0, y_0)(x_1 - x_0).$$

На участке $[x_1, x_2]$ заменим интегральную кривую отрезком M_1M_2 прямой, выходящей из точки $M_1(x_1, y_1)$ в направлении поля в точке M_1 . Этот отрезок прямой будет иметь угловой коэффициент

$$\operatorname{tg} \alpha_1 = y'|_{M_1} = f(x_1, y_1),$$

а координата y_2 точки $M_2(x_2, y_2)$ определится по формуле

$$y_2 = y_1 + f(x_1, y_1)(x_2 - x_1).$$

На участке $[x_2, x_3]$ интегральную кривую заменим отрезком прямой M_2M_3 , выходящим из точки M_2 в направлении поля, и т. д. Указанный процесс можно продолжать до тех пор, пока мы не выйдем из области определения функции $f(x, y)$.

Построенная таким образом ломаная линия $M_0M_1M_2\dots$, называемая *ломаной Эйлера*, представляет приближенное изображение интегральной кривой уравнения (1), проходящей через точку (x_0, y_0) .

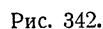
Примечание. Построение ломаной Эйлера можно несколько упростить. Возьмем на оси OX слева от начала координат полюс P на расстоянии $PO = 1$ и будем откладывать на оси OY последовательно отрезки OA_1, OA_2, OA_3 и т. д. так, чтобы $\overline{OA_1} = f(x_0, y_0)$, $\overline{OA_2} = f(x_1, y_1)$, $\overline{OA_3} = f(x_2, y_2)$ и т. д.

Прямая PA_1 имеет угловой коэффициент $\frac{\overline{OA_1}}{1} = f(x_0, y_0)$, т. е. звено M_0M_1 ломаной проводится параллельно прямой PA_1 . Прямая PA_2 имеет угловой коэффициент $f(x_1, y_1)$, т. е. звено M_1M_2 ломаной проводится параллельно прямой PA_2 , и т. д. (рис. 342).

При построении отрезков PO и $OA_1, OA_2, \dots$ можно пользоваться масштабом, отличным от принятого на осях координат.

Ломаная $M_0M_1M_2\dots$, вообще говоря, имеет с искомой интегральной кривой лишь одну общую точку $M_0(x_0, y_0)$ и по мере удаления от этой точки отклонение ломаной от интегральной кривой может возра-

Вычисление ординат $y_1, y_2, \dots, y_n$ вершин ломаной Эйлера можно, не строя самой ломаной, производить последовательно по формулам:



[illegible]

$$x_1 = x_0 + h, \quad x_2 = x_0 + 2h, \quad \dots, \quad x_n = x_0 + nh,$$

Сами вычисления удобно располагать в виде таблицы, заполняя последовательно одну строку за другой.

x	y	$f(x, y)$	$f(x, y) \hbar$
x_0	y_0	$f(x_0, y_0)$	$f(x_0, y_0) \hbar$
x_1	y_1	$f(x_1, y_1)$	$f(x_1, y_1) \hbar$
x_2	y_2	$f(x_2, y_2)$	$f(x_2, y_2) \hbar$
x_3	y_3		

Пример. Чтобы иметь возможность сравнить приближенный результат с точным, рассмотрим уравнение $y' = x + y$, интегрируемое в квадратурах.

Пусть требуется определить численное значение решения этого уравнения с начальными условиями 0;1 при $x=0,5$.

Интегрируя данное линейное уравнение методом вариации постоянных, получим:

$$\frac{dy}{dx} = y, \quad \frac{dy}{y} = dx, \quad \ln |y| = x + \ln |C|, \quad y = Ce^x$$

— решение соответствующего однородного уравнения,

$$\begin{aligned} y &= C(x) e^x, \quad y' = C'(x) e^x + C(x) e^x, \\ C'(x) e^x + C(x) e^x &= C(x) e^x + x, \quad C'(x) = x e^{-x}, \\ C(x) &= -x e^{-x} - e^{-x} + C, \quad y = C e^x - x - 1, \end{aligned}$$

где C — произвольная постоянная, — общее решение данного уравнения.

Используя начальные условия 0;1, находим $C=2$. Значение решения $y=2e^x - x - 1$ при $x=0,5$ вычисляем приближенно: $y(0,5) \approx 1,797$.

Проведем теперь вычисление значения $y(0,5)$ методом Эйлера. Для этого разобьем отрезок $[0, 0,5]$ точками $x_0=0, x_1=0,1, x_2=0,2, x_3=0,3, x_4=0,4, x_5=0,5$ на пять равных частей (шаг вычисления $h=0,1$). Обозначая через $y_0, y_1, y_2, y_3, y_4, y_5$ ординаты вершин соответствующей ломаной Эйлера, получим таблицу:

x	y	$f(x, y) = x + y$	$f(x, y) \cdot h = (x + y) 0,1$
0	1	1	0,1
0,1	1,1	1,2	0,12
0,2	1,22	1,42	1,142
0,3	1,362	1,662	1,166
0,4	1,528	1,928	0,193
0,5	1,721		

Найденное приближенное численное значение решения $y(0,5)$ равно 1,721. Ошибка не превосходит 0,08.

Практическая погрешность от вычисления численных значений решения уравнения $y' = f(x, y)$ методом Эйлера складывается из погрешности самого метода Эйлера, заменяющего точные значения решения приближенными, а также из погрешностей, допускаемых при приближенных вычислениях значений функции $f(x, y)$ в вершинах ломаной и при приближенных вычислениях, связанных с округлениями чисел $y_1, y_2, \dots, y_n$. Не имея возможности подробно останавливаться на этих вопросах, отсылаем лиц, занимающихся приближенными вычислениями, к специальной литературе.

Следует отметить, что степень точности самого метода Эйлера вообще весьма невелика, в связи с чем для отыскания этим методом решения с наперед заданной степенью точности обычно приходится применять так называемые *формулы пересчета* каждого из значений $y_1, y_2, \dots, y_n$.

Найденное значение y_1 уточняется тем, что в формуле $y_1 = y_0 + f(x_0, y_0)h$ вместо $y_0 = f(x_0, y_0)$ берется среднее арифметическое значений $f(x_0, y_0)$ и $f(x_1, y_1)$. В качестве уточненного значения y_1 берется значение

$$y_1^{(1)} = y_0 + \frac{f(x_0, y_0) + f(x_1, y_1)}{2} h,$$

затем находится значение производной y' в точке $(x_1, y_1^{(1)})$. Заменяя $f(x_0, y_0)$ средним арифметическим значений $f(x_0, y_0)$ и $f(x_1, y_1^{(1)})$, находят второе уточненное значение y_1 :

$$y_1^{(2)} = y_0 + \frac{f(x_0, y_0) + f(x_1, y_1^{(1)})}{2} h,$$

затем третье:

$$y_1^{(3)} = y_0 + \frac{f(x_0, y_0) + f(x_1, y_1^{(2)})}{2} h$$

и т. д., пока два последовательных уточненных значения y_1 не совпадут между собой в пределах заданной степени точности. Эти совпадающие значения $y_1^{(i)}$ и $y_1^{(i+1)}$ принимаются за ординату вершины M_1 уточненной ломаной Эйлера, соответствующей значению x_1 .

Аналогичным способом уточняют значения y для $x_2, x_3, \dots$. Так, например, за $y_2^{(1)}$ принимают

$$y_2^{(1)} = y_1 + f(x_1, y_1) h,$$

где y_1 — найденное уточненное значение ординаты вершины ломаной с абсциссой x_1 .

За второе уточнение y_2 принимают

$$y_2^{(2)} = y_1 + \frac{f(x_1, y_1) + f(x_2, y_2^{(1)})}{2} h,$$

за третье —

$$y_2^{(3)} = y_1 + \frac{f(x_1, y_1) + f(x_2, y_2^{(2)})}{2} h$$

и т. д., пока приближения y_2 не совпадут в пределах заданной степени точности.

Если уточнение ординаты не приводит к цели, после нескольких уточнений уменьшают шаг h .

Вернемся к рассмотренному выше примеру.

Применяя формулы пересчета значений, получим:

$$\begin{aligned} y_1^{(1)} &= y_1^{(2)} = 1,100 = y_1, \\ y_2^{(1)} &= y_2^{(2)} = 1,243 = y_2, \\ y_3^{(2)} &= y_3^{(3)} = 1,400 = y_3, \\ y_4^{(2)} &= y_4^{(3)} = 1,585 = y_4, \\ y_5^{(2)} &= y_5^{(3)} = 1,799 = y_5. \end{aligned}$$

Найденное значение y_5 отличается от значения $y(0,5) \approx 1,797$ не более чем на 0,002.

3. Метод Адамса — Крылова. При построении ломаной Эйлера уравнения $y' = f(x, y)$, соответствующей начальной точке (x_0, y_0) и системе значений $x_0, x_1, x_2, \dots, x_n$, кривая на каждом отрезке деления оси OX заменялась прямой, угловым коэффициент которой в одной из концевых точек отрезка совпадал с угловым коэффициентом касательной к интегральной кривой, проходящей через эту точку. Замена интегральной кривой такой ломаной является довольно грубым приближением решения. Можно усовершенствовать этот метод, заменяя на отрезках деления решение не линейными функциями, а многочленами более высокой степени.

Пусть требуется найти решение уравнения

$$y' = f(x, y) \quad (1)$$

с начальными данными x_0, y_0 .

Рассмотрим систему значений переменной x : $x_0, x_1, x_2, \dots, x_k, \dots$ ($x_0 < x_1 < x_2 < \dots < x_k < \dots$)¹⁾ и вычислим каким-нибудь способом (методом последовательных приближений, методом Эйлера или с помощью

¹⁾ Для убывающей системы значений x рассуждения производятся аналогичным образом.

Вывод этих формул основывается на использовании так называемой второй интерполяционной формулы Ньютона.

Рассмотрим систему факториальных многочленов, определяемых равенствами:

Очевидно, что $\Phi_{ik}(x)$ есть многочлен i -й степени, причем в любой точке x

Пусть $P(x)$ — многочлен степени n , принимающий в узлах интерполяции $x_0, x_1, \dots, x_n$ заданные значения $y_0, y_1, \dots, y_n$:

Обозначим через $C_0, C_1, \dots, C_n$ коэффициенты разложения многочлена $P(x)$ по факториальным многочленам $\Phi_{0n}(x), \Phi_{1n}(x), \dots, \Phi_{nn}(x)$, имеющим соответственно степени $0, 1, 2, \dots, n$.

$$\begin{array}{lcl} P(x) = & C_0 + C_1 \Phi_{1n}(x) + C_2 \Phi_{2n}(x) + \dots + C_n \Phi_{nn}(x), \\ \Delta P(x) = & C_1 & + C_2 \Phi_{1n-1}(x) + \dots + C_n \Phi_{n-1\ n-1}(x), \\ \Delta^2 P(x) = & & C_2 & + \dots + C_n \Phi_{n-2\ n-2}(x), \\ & \dots & & \\ \Delta^n P(x) = & & & C_n. \end{array}$$
$$P(x_n) = C_0, \quad \Delta P(x_{n-1}) = C_1, \quad \Delta^2 P(x_{n-2}) = C_2, \dots, \quad \Delta^n P(x_0) = C_n$$
$$y_n = C_0, \quad \Delta y_{n-1} = C_1, \quad \Delta^2 y_{n-2} = C_2, \dots, \quad \Delta^n y_0 = C_n.$$
$$P(x) = y_n + \sum_{i=1}^n \Delta^i y_{n-i} \frac{(x-x_n)(x-x_{n-1}) \dots (x-x_{n-i+1})}{i! h^i}.$$

Пусть $[x_n, x_{n+1}]$ — любой отрезок с номером $n \geq k$, а $r(x)$ — соответствующий ему многочлен степени не выше k , фигурирующий в определении ломаной Эйлера $y = L(x)$. Обозначим

[по определению многочлена $r(x)$]

и запишем разложение многочлена $r(x)$ по факториальным многочленам с помощью второй интерполяционной формулы Ньютона:

¹⁾ Здесь $\Delta\Phi_{ik}(x) = \Phi_{ik}(x + \hbar) - \Phi_{ik}(x)$. Справедливость данного равенства устанавливается непосредственным подсчетом разности, как это делалось в гл. X, § 29.

Интегрируя уравнение $y' = r(x)$ на участке $[x_n, x_{n+1}]$, получим:

$$\begin{aligned} \Delta y_n &= y_{n+1} - y_n = \int_{x_n}^{x_{n+1}} r(x) dx = \\ &= \int_{x_n}^{x_n+h} \left[y'_n + \sum_{i=1}^k \Delta^i y'_{n-i} \frac{(x-x_n)(x-x_{n-1}) \dots (x-x_{n-i+1})}{i! h^i} \right] dx = \\ &= y'_n h + \sum_{i=1}^k \frac{\Delta^i y'_{n-i}}{i! h^i} \int_{x_n}^{x_n+h} (x-x_n)(x-x_{n-1}) \dots (x-x_{n-i+1}) dx = \\ &= y'_n h + \sum_{i=1}^k \frac{\Delta^i y'_{n-i}}{i!} h \int_0^1 t(t+1) \dots (t+i-1) dt = \left(y'_n + \sum_{i=1}^k a_i \Delta^i y'_{n-i} \right) h, \end{aligned}$$

где

$$a_i = \frac{1}{i!} \int_0^1 t(t+1) \dots (t+i-1) dt.$$

Таким образом, выведена формула

$$y_{n+1} = y_n + h \left(y'_n + \sum_{i=1}^k a_i \Delta^i y'_{n-i} \right),$$

позволяющая последовательно определять значения $y_{k+1}, y_{k+2}, \dots$ для вершин ломаной Эйлера порядка $k+1$. При $k=3$ имеем

$$y_{n+1} = y_n + h \left[y'_{n-1} + \frac{1}{2} \Delta y'_{n-1} + \frac{5}{12} \Delta^2 y'_{n-2} + \frac{3}{8} \Delta^3 y'_{n-3} \right].$$

Предложенный метод графического и численного интегрирования уравнения $y' = f(x, y)$ с начальными данными (x_0, y_0) называется *методом Адамса—Крылова*. При $k=0$ из него как частный случай получается рассмотренный выше метод Эйлера.

Оценка погрешности при применении метода Адамса—Крылова довольно сложна, в связи с чем для получения нужной степени точности вычислений используется ряд специальных практических приемов, которые в данном курсе не рассматриваются.

Найдем, в качестве примера, методом Адамса—Крылова значение решения уравнения $y' = x + y$ с начальными данными $0;1$ при $x=0,5$, взяв шаг вычисления $h=0,1$.

Значения решения в точках $x_0=0,0, x_1=0,1, x_2=0,2, x_3=0,3, x_4=0,4$ определим методом Эйлера (см. п. 2):

$$y_0 = 1,000, y_1 = 1,110, y_2 = 1,243, y_3 = 1,400, y_4 = 1,585.$$

Значение $y_5 \approx y(0,5)$ найдем по формуле Адамса—Крылова (для $k=4$):

$$y_5 = y_4 + h [y'_4 + a_1 \Delta y'_3 + a_2 \Delta^2 y'_2 + a_3 \Delta^3 y'_1 + a_4 \Delta^4 y'_0],$$

где

$$\begin{aligned} a_1 &= \int_0^1 t dt = \frac{1}{2}, \quad a_2 = \frac{1}{2!} \int_0^1 t(t+1) dt = \frac{5}{12}, \\ a_3 &= \frac{1}{3!} \int_0^1 t(t+1)(t+2) dt = \frac{3}{8}, \quad a_4 = \frac{1}{4!} \int_0^1 t(t+1)(t+2)(t+3) dt = \frac{251}{720}. \end{aligned}$$

Вычисления удобно располагать в виде следующей таблицы ¹⁾:

x	y	$y' = x + y$	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$
0,0	1,000	1,000				
			210			
0,1	1,110	1,210		+ 23		
			233		+ 1	
0,2	1,243	1,443		+ 24		+ 3
			257	-	+ 4	
0,3	1,400	1,700		+ 28		
			285			
0,4	1,585	1,985				
0,5						

При этом

$$y_5 = 1,585 + 0,1 \left[1,985 + \frac{1}{2} \cdot 0,285 + \frac{5}{12} \cdot 0,028 + \frac{3}{8} \cdot 0,004 + \frac{251}{720} \cdot 0,003 \right] \approx 1,799.$$

Если потребуется продолжить вычисления для $x_6 = 0,6$, то надо внести в таблицу значение y_5 , дополнить остальные столбцы и определить y_6 по формуле

$$y_6 = y_5 + h \left[y'_5 + \frac{1}{2} \Delta y'_5 + \frac{5}{12} \Delta^2 y'_5 + \frac{3}{8} \Delta^3 y'_5 + \frac{251}{720} \Delta^4 y'_5 \right]$$

и т. д.

§ 9. Применение дифференциальных уравнений первого порядка к решению некоторых задач механики и физики

Решение большого количества математических и физических задач приводит к составлению и последующему интегрированию дифференциальных уравнений первого порядка. Составление дифференциальных уравнений, особенно в задачах прикладного характера, обычно вызывает затруднения, так как для этого необходимо знание разнообразных физических законов и наличие известных навыков в переводе физических задач на математический язык. Универсального рецепта для составления дифференциальных уравнений не существует. В одних случаях полезно (особенно в задачах геометрического характера) использовать геометрический смысл производной от функции как тангенса угла наклона касательной к графику функции, в других — механический смысл первой и второй производных как скорости и ускорения процесса, иногда целесо-

¹⁾ Все табличные разности вписываются в таблицу в целых единицах младшего разряда чисел y' .

образнее оказывается составить приближенные соотношения между бесконечно малыми приращениями искомых величин или дифференциалами этих величин и т. д.

Рассмотрим несколько задач на составление и интегрирование дифференциальных уравнений первого порядка.

1. Задача об охлаждении тела. Ньютон установил, что скорость изменения температуры охлаждающегося тела пропорциональна разности температур тела и окружающей среды (температура среды считается постоянной).

Обозначим через T температуру тела в момент времени t , а через $T_{\text{ср.}}$ — температуру окружающей среды ($T_{\text{ср.}} < T$). Скорость изменения температуры тела дается производной от температуры по времени. Следовательно, по закону Ньютона:

$$\frac{dT}{dt} = -k(T - T_{\text{ср.}}),$$

где k — коэффициент пропорциональности, $k > 0$.

Это дифференциальное уравнение с разделяющимися переменными. Интегрируя, получаем:

$$T = T_{\text{ср.}} + Ce^{-kt}.$$

Для определения постоянной C необходимо задание начальных условий процесса. Пусть, например, в момент времени t_0 температура тела равнялась T_0 . Тогда

$$T_0 = T_{\text{ср.}} + C_0 e^{-kt_0}, \quad C_0 = (T_0 - T_{\text{ср.}}) e^{kt_0},$$

т. е. закон изменения температуры охлаждающегося тела имеет вид:

$$T = T_{\text{ср.}} + (T_0 - T_{\text{ср.}}) e^{-k(t-t_0)}.$$

Коэффициент k либо задается непосредственно, либо определяется с помощью дополнительного условия:

$$T|_{t=t_1} = T_1.$$

2. Барометрическая формула. Найдем зависимость между высотой h над уровнем моря и давлением воздуха p .

Давление воздуха на горизонтальную площадку в 1 см^2 определяется весом призматического столба воздуха, расположенного над этой площадкой. Пусть на высоте h над уровнем моря давление воздуха на горизонтальную площадку равно p . На высоте $h + \Delta h$, $\Delta h > 0$, давление воздуха равно $p + \Delta p$, где $\Delta p < 0$. Уменьшение давления численно равно весу столба воздуха, заключенного между площадками. Так как по закону Бойля — Мариотта плотность воздуха при постоянной температуре пропорциональна давлению, то на высоте h она равна kp , где k — коэффициент пропорциональности, $k > 0$, а на высоте $h + \Delta h$ равна $k(p + \Delta p)$. Считая плотность внутри столба высотой Δh постоянной и равной некоторому «среднему» значению плотности, т. е. $k(p + \theta \Delta p)$, $0 < \theta < 1$, получаем соотношение

$$\Delta p = -k(p + \theta \Delta p) \Delta h,$$

точность которого повышается с уменьшением Δh ¹⁾. Разделив обе части этого соотношения на Δh и перейдя к пределу при стремлении Δh к нулю, получим дифференциальное уравнение, связывающее величины h и p :

$$\frac{dp}{dh} = -kp.$$

Решения этого уравнения имеют вид: $p = Ce^{-kh}$. Принимая давление на уровне моря ($h=0$) равным p_0 , запишем барометрическую формулу в виде: $p = p_0 e^{-kh}$ или, разрешая относительно h , в виде: $h = \frac{1}{k} \ln \frac{p_0}{p}$.

Эта формула позволяет судить о высоте места над уровнем моря, если известно давление p .

3. Задача о прожекторе. Определить форму зеркала, отражающего все лучи, выходящие из данной точки O параллельно данному направлению $\vec{l}$.

Искомая поверхность есть поверхность вращения. Рассмотрим любую плоскость, проходящую через точку O параллельно $\vec{l}$. Введем на ней систему координат, приняв точку O за начало координат и направив ось OX по данному направлению $\vec{l}$ (рис. 343). Поверхность зеркала пересекает выбранную плоскость по кривой, уравнение которой $y = f(x)$. Пусть $M(x, y)$ — любая точка этой кривой, ML — касательная, MQ — нор-

¹⁾ Предполагается, что p непрерывно зависит от h .

маль к кривой в этой точке. Так как угол падения луча равен углу отражения, имеем $\angle OMQ = \angle QMN$. Следовательно,

$$\angle OLM = \angle OML, \quad OL = OM = \sqrt{x^2 + y^2}, \quad \overline{LP} = x + \sqrt{x^2 + y^2}, \quad \overline{OE} = y, \\ \operatorname{tg} \angle MLP = \frac{y}{x + \sqrt{x^2 + y^2}}.$$

Учитывая геометрический смысл производной y' , получаем дифференциальное уравнение для определения искомой кривой:

$$y' = \frac{y}{x + \sqrt{x^2 + y^2}}.$$

Это — однородное дифференциальное уравнение первого порядка. Выбирая x за неизвестную функцию, а y за аргумент, перепишем полученное уравнение в виде:

$$\frac{dx}{dy} = \frac{x}{y} + \sqrt{\frac{x^2}{y^2} + 1}.$$

Для решения уравнения применим подстановку $x = uy$, где u — новая неизвестная функция y . При этом

$$\frac{dx}{dy} = u + \frac{du}{dy} y, \\ u + \frac{du}{dy} y = u + \sqrt{u^2 + 1}, \\ \frac{du}{\sqrt{u^2 + 1}} = \frac{dy}{y},$$

$$\ln |u + \sqrt{u^2 + 1}| = \ln |y| + \ln |C|,$$

$u + \sqrt{u^2 + 1} = Cy$ — общий интеграл вспомогательного уравнения.

$$\text{Так как } u + \sqrt{u^2 + 1} = Cy, \text{ то } \frac{1}{u + \sqrt{u^2 + 1}} = \frac{1}{Cy}, \text{ т. е. } \frac{u - \sqrt{u^2 + 1}}{-1} = \frac{1}{Cy}.$$

Складывая соотношения $u + \sqrt{u^2 + 1} = Cy$ и $u - \sqrt{u^2 + 1} = -\frac{1}{Cy}$, находим общее решение вспомогательного уравнения в виде:

$$u = \frac{1}{2} \left(Cy - \frac{1}{Cy} \right).$$

Общее решение данного уравнения получим, заменив функцию u ее выражением через x и y :

$$x = \frac{1}{2} \left(Cy^2 - \frac{1}{C} \right), \quad C^2 y^2 = 2Cx + 1.$$

Таким образом, искомая кривая — парабола, в фокусе которой помещается источник света, а искомая отражательная поверхность — параболоид вращения.

4. Задача о колебаниях в электрической цепи. Электродвижущая сила E включается в некоторый момент времени ($t=0$) в контур, состоящий из последовательно соединенных катушки самоиндукции L и сопротивления R . Требуется найти ток I в момент времени t , если в начальный момент времени он был равен нулю.

По закону Кирхгофа электродвижущая сила в цепи равна сумме падений напряжения на индуктивном и активном сопротивлениях:

$$E = u_L + u_R,$$

где

$$u_L = L \frac{dI}{dt}, \quad u_R = RI.$$

Таким образом,

$$L \frac{dI}{dt} + RI = E.$$

Предположим, что электродвижущая сила меняется по закону: $E = B \sin \omega t$. Задача определения закона изменения тока в цепи сводится, таким образом, к решению линейного неоднородного уравнения:

$$\frac{dI}{dt} + \frac{R}{L} I = \frac{B}{L} \sin \omega t$$

с начальными данными $t_0 = 0, I_0 = 0$.

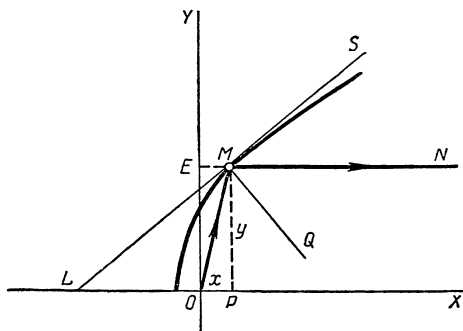


Рис. 343.

Решаем сначала однородное уравнение

$$\frac{dI}{dt} + \frac{R}{L} I = 0.$$

Его общее решение:

$$I = C e^{-\frac{R}{L} t}.$$

Ищем решение неоднородного уравнения в виде:

$$I = C(t) e^{-\frac{R}{L} t}.$$

Подставляя $I = C(t) e^{-\frac{R}{L} t}$ и $I' = C'(t) e^{-\frac{R}{L} t} - \frac{R}{L} C(t) e^{-\frac{R}{L} t}$ в неоднородное уравнение, после упрощений получаем:

$$C'(t) = \frac{B}{L} e^{\frac{R}{L} t} \sin \omega t,$$

откуда

$$C(t) = \frac{B e^{\frac{R}{L} t} (R \sin \omega t - L \omega \cos \omega t)}{R^2 + \omega^2 L^2} + C.$$

Таким образом,

$$I(t) = C e^{-\frac{R}{L} t} + \frac{B (R \sin \omega t - L \omega \cos \omega t)}{R^2 + \omega^2 L^2}.$$

Постоянную C_0 определяем с помощью начальных условий:

$$0 = C_0 - \frac{BL\omega}{R^2 + \omega^2 L^2}, \quad C_0 = \frac{BL\omega}{R^2 + \omega^2 L^2}.$$

При этом

$$I(t) = \frac{BL\omega}{R^2 + \omega^2 L^2} e^{-\frac{R}{L} t} + \frac{B}{R^2 + \omega^2 L^2} (R \sin \omega t - L \omega \cos \omega t).$$

Если обозначить

$$\frac{R}{\sqrt{R^2 + \omega^2 L^2}} = \cos \alpha, \quad -\frac{L\omega}{\sqrt{R^2 + \omega^2 L^2}} = \sin \alpha,$$

то закон изменения силы тока запишется в виде:

$$I(t) = \frac{BL\omega}{R^2 + \omega^2 L^2} e^{-\frac{R}{L} t} + \frac{B}{\sqrt{R^2 + \omega^2 L^2}} \sin(\omega t + \alpha).$$

Отсюда следует, что величину тока в цепи можно рассматривать как сумму двух слагаемых, одно из которых носит затухающий характер, а второе периодический. При больших значениях t ток в цепи фактически определяется вторым слагаемым, т. е. в цепи будут совершаться электрические колебания.

Раздел 2

ОБЩИЕ СВЕДЕНИЯ О ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЯХ ВЫСШЕГО ПОРЯДКА

§ 10. Дифференциальные уравнения высшего порядка. Общие понятия

Как уже говорилось выше, *дифференциальным уравнением n -го порядка* называется соотношение

$$F(x, y, y', \dots, y^{(n)}) = 0, \quad (1)$$

связывающее независимую переменную, искомую функцию и ее производные до n -го порядка включительно. Всякое дифференциальное уравнение порядка выше первого называется уравнением *высшего порядка*.

Так, например, уравнения $y''' - 1 = 0$, $xy' - 2y'' + y'^3 = 0$, $y^{v^3} - e^{iy'''} + 1 = 0$ — уравнения высших (соответственно третьего, второго и пятого) порядков.

В некоторых случаях уравнение (1) удастся разрешить относительно $y^{(n)}$, т. е. записать в виде:

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}). \quad (2)$$

Такое уравнение называется уравнением n -го порядка, *разрешенным относительно старшей производной*.

Так же как и дифференциальные уравнения первого порядка, дифференциальные уравнения высших порядков имеют, вообще говоря, бесчисленные множества решений, каждое из этих решений изображается на плоскости XOY некоторой кривой (интегральной кривой).

Дифференциальному уравнению второго порядка

$$F(x, y, y', y'') = 0 \quad (3)$$

легко дать геометрическое и механическое истолкования.

Кривизна кривой $y = \varphi(x)$ в каждой ее точке вычисляется по формуле

$$K = \frac{y''}{(1 + y'^2)^{3/2}}.$$

Записав уравнение (3) в виде:

$$F\left(x, y, y', \frac{y''}{(1 + y'^2)^{3/2}} (1 + y'^2)^{3/2}\right) = 0,$$

видим, что дифференциальное уравнение второго порядка выражает зависимость между координатами точки интегральной кривой, угловым коэффициентом ее касательной и кривизной в этой точке. Интегральные кривые уравнения (3) — это кривые, которые в каждой своей точке имеют предписываемое уравнением соотношение между угловым коэффициентом касательной к кривой и кривизной.

Если независимую переменную (обозначим ее через t) рассматривать как время, а искомую функцию y — как путь, пройденный прямолинейно движущейся точкой за время t , то дифференциальное уравнение

$$F\left(t, y, \frac{dy}{dt}, \frac{d^2y}{dt^2}\right) = 0 \quad (4)$$

выражает в каждый момент времени t зависимость между пройденным путем y , скоростью $\frac{dy}{dt}$ и ускорением $\frac{d^2y}{dt^2}$ движущейся точки. Решить такое уравнение — значит определить по указанной зависимости закон движения, т. е. дать соотношение $y = \varphi(t)$, позволяющее в любой момент времени t определять положение движущейся точки. Уравнение (4) определяет, вообще говоря, бесчисленное множество решений (их называют *движениями*). Для того чтобы из этого бесчисленного множества движений выбрать определенное, в механике обычно задают начальное положение точки, т. е. значение y при $t = t_0$ (обозначим его y_0), и начальную скорость, т. е. значение $\frac{dy}{dt}$ при $t = t_0$ (обозначим ее y'_0).

Так, например, при решении задачи о притяжении материальной точки m неподвижным центром O (см. § 1) мы получили бесчисленное множество решений уравнения $my'' = -ky$ (k — коэффициент пропорциональности) в виде:

$$y = C_1 \cos \omega t + C_2 \sin \omega t,$$

где $\omega = \sqrt{\frac{k}{m}}$, а C_1 и C_2 — произвольные постоянные.

Если в начальный момент времени ($t_0=0$) точка m находилась от центра притяжения на расстоянии a_0 ($y_0=a_0$) и не имела начальной скорости ($y'_0=0$), то постоянные C_1 и C_2 однозначно определяются из соотношений

$$\begin{aligned} y &= C_1 \cos \omega t + C_2 \sin \omega t, & C_1 &= a_0, \\ y' &= -C_1 \omega \sin \omega t + C_2 \omega \cos \omega t, & C_2 \omega &= 0, \quad C_2 = 0, \end{aligned}$$

и мы получаем конкретное колебательное движение:

$$y = a_0 \cos \omega t.$$

В связи с тем что дифференциальные уравнения широко применяются в механике, независимо от конкретного физического или геометрического смысла аргумента x и искомой функции y , числа x_0 , y_0 , y'_0 , представляющие собой некоторое значение аргумента ($x=x_0$) и значения искомой функции ($y=y_0$) и ее производной ($y'=y'_0$) при этом значении аргумента принято называть *начальными условиями* или *начальными данными* для уравнения второго порядка:

$$F(x, y, y', y'') = 0.$$

Говорят, что решение $y=\varphi(x)$ этого уравнения удовлетворяет начальным условиям x_0 , y_0 , y'_0 , если $\varphi(x_0)=y_0$, $\varphi'(x_0)=y'_0$.

Геометрически это значит, что соответствующая интегральная кривая уравнения проходит через точку (x_0, y_0) плоскости $ХОУ$ и имеет в этой точке касательную с угловым коэффициентом y'_0 .

Не останавливаясь на механической и геометрической интерпретации дифференциальных уравнений третьего и более высоких порядков, заметим, что для отыскания конкретного решения уравнения третьего порядка

$$F(x, y, y', y'', y''') = 0$$

достаточно, вообще говоря, задать значения y , y' , y'' при некотором значении $x=x_0$ (начальные условия x_0 , y_0 , y'_0 , y''_0), для уравнения четвертого порядка — значения y , y' , y'' , y''' при $x=x_0$ (начальные условия x_0 , y_0 , y'_0 , y''_0 , y'''_0) и т. д.

Для уравнения n -го порядка (1) или (2) *начальными условиями* называют систему из $(n+1)$ -го чисел x_0 , y_0 , y'_0 , ..., $y^{(n-1)}_0$, представляющих начальное значение независимой переменной x ($x=x_0$) и значения искомой функции y и всех ее производных до $(n-1)$ -го порядка включительно при $x=x_0$.

Решение $y=\varphi(x)$ уравнения (1) или (2) удовлетворяет начальным условиям x_0 , y_0 , y'_0 , ..., $y^{(n-1)}_0$, если

$$\varphi(x_0)=y_0, \quad \varphi'(x_0)=y'_0, \quad \dots, \quad \varphi^{(n-1)}(x_0)=y^{(n-1)}_0.$$

Отыскание решения уравнения (1) или (2), удовлетворяющего заданному начальным условиям x_0 , y_0 , y'_0 , ..., $y^{(n-1)}_0$, называется решением *задачи Коши* для этих уравнений.

Достаточные условия существования и единственности решения задачи Коши для уравнения, разрешенного относительно старшей производной, даются теоремой, сформулированной по аналогии с теоремой Коши для уравнения первого порядка на геометрическом языке.

Теорема Коши. Если функция $(n+1)$ -го переменного $f(x, y, y', \dots, y^{(n-1)})$ в некоторой области D $(n+1)$ -мерного пространства непрерывна и имеет непрерывные частные производные по y , y' , ..., $y^{(n-1)}$, то, какова бы ни была точка $(x_0, y_0, y'_0, \dots, y^{(n-1)}_0)$ этой области, существует, и притом единственное, решение $y=\varphi(x)$ уравнения

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}),$$

определенное в некотором интервале, содержащем точку x_0 , удовлетворяющее начальным условиям x_0 , y_0 , y'_0 , ..., $y^{(n-1)}_0$.

Заметим, что единственность решения задачи Коши для уравнения n -го порядка ($n \geq 1$) не означает, что через данную точку (x_0, y_0)

Например, для уравнения второго порядка с начальными условиями x_0, y_0, y'_0 единственность решения задачи Коши нужно понимать в том смысле, что через точку (x_0, y_0) плоскости XOY проходит единственная интегральная кривая уравнения, касательная к которой в этой точке имеет угловой коэффициент y'_0 . Через ту же точку (x_0, y_0) проходит еще бесчисленное множество интегральных кривых уравнения с другим наклоном касательной в этой точке.

$$y = \varphi(x, y_0, y'_0, \dots, y_0^{(n-1)}).$$

К тому же выводу приводит решение задачи о составлении дифференциального уравнения семейства кривых, зависящего от n параметров, т. е. задачи, в некотором смысле обратной задаче интегрирования дифференциального уравнения.

$$\Phi(x, y, C_1, C_2, \dots, C_n) = 0, \quad (5)$$

Так, например, семейство всех невертикальных прямых плоскости XOY имеет уравнение $y = C_1x + C_2$, C_1 и C_2 — параметры. Каждой прямой этого семейства соответствуют определенные значения параметров C_1 и C_2 .

Пусть требуется составить дифференциальное уравнение, для которого кривые семейства (5) будут являться интегральными кривыми.

$$\Phi(x, y, C_1, C_2, \dots, C_n) = 0,$$

$$\frac{d\Phi}{dx} = \Phi'_x + \Phi'_y \cdot y' = 0,$$

$$\frac{d^2\Phi}{dx^2} = \Phi''_{xx} + \Phi''_{xy}y' + \Phi''_{yx}y' + \Phi''_{yy}y'^2 + \Phi'_y y'' = 0,$$

$$\frac{d^n \Phi}{d x^n} = \Phi_{xx \dots x}^{(n)} + \dots + \Phi'_y \cdot y^{(n)} = 0.$$

$$F(x, y, y', \dots, y^{(n)}) = 0. \quad (6)$$

Так как кривые семейства (5) различаются значениями параметров $C_1, C_2, \dots, C_n$, а полученное соотношение от этих параметров не зависит, то, рассуждая аналогичным образом для любой другой кривой семейства (5), снова приходим к тому же соотношению.

Это соотношение, являющееся дифференциальным уравнением n -го порядка, называется *дифференциальным уравнением семейства* (5). Оно выражает общее свойство всех кривых семейства. Любая из кривых семейства (5), зависящего от n параметров, является интегральной кривой этого уравнения.

Составим, например, дифференциальное уравнение семейства окружностей радиуса R :

$$(x - C_1)^2 + (y - C_2)^2 = R^2, \quad (7)$$

зависящего от двух параметров C_1 и C_2 .

Для этого два раза последовательно дифференцируем по x уравнение семейства, считая параметры C_1 и C_2 постоянными, а y — неявной функцией x .

Получаются соотношения:

$$\begin{aligned} 2(x - C_1) + 2(y - C_2) \cdot y' &= 0, \\ 1 + y''(y - C_2) + y'^2 &= 0. \end{aligned} \quad (8)$$

Исключая из соотношений (7) и (8) параметры C_1 и C_2 , находим:

$$\begin{aligned} y - C_2 &= -\frac{1 + y'^2}{y''}, \quad x - C_1 = \frac{(1 + y'^2)y'}{y''}, \\ \frac{(1 + y'^2)^2 y'^2}{y''^3} + \frac{(1 + y'^2)^2}{y''^3} &= R^2, \end{aligned}$$

т. е. $(1 + y'^2)^3 = R^2 y''^3$ — дифференциальное уравнение данного семейства окружностей.

Дадим теперь определение общего и частного решений уравнения n -го порядка, разрешенного относительно старшей производной.

Функция $y = \varphi(x, C_1, C_2, \dots, C_n)$, зависящая от n произвольных постоянных $C_1, C_2, \dots, C_n$, называется *общим решением* уравнения $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$ в некоторой области σ плоскости $ХОУ$, если она является решением этого уравнения для любых значений постоянных $C_1, C_2, \dots, C_n$ (или хотя бы любых значений этих постоянных из некоторого множества) и если любое решение уравнения, лежащее в области σ , может быть записано в виде $y = \varphi(x, C_1, C_2, \dots, C_n)$ при конкретных значениях $C_1, C_2, \dots, C_n$.

Решения, получающиеся из общего при конкретных значениях постоянных $C_1, C_2, \dots, C_n$, называются *частными решениями* данного уравнения.

Неявно заданное общее или частное решение уравнения называются соответственно его *общим* и *частным интегралами*.

Геометрически общее решение (общий интеграл) уравнения (2) в области σ представляет собой семейство интегральных кривых, зависящее от n параметров, обладающее тем свойством, что каждая интегральная кривая, лежащая в области σ , принадлежит этому семейству кривых.

При определенных условиях (см., например, условия теоремы Коши), каждой системе начальных данных $x_0, y_0, y'_0, \dots, y_0^{(n-1)}$, где (x_0, y_0) — любая точка области σ , соответствует, и притом одно, решение уравнения (2), лежащее в области σ . Если функция $y = \varphi(x, C_1, C_2, \dots, C_n)$ — общее решение уравнения (2) в области σ , то существуют значения $C_1, C_2, \dots, C_n$, соответствующие этому решению, а значит, решение, определяемое начальными условиями $x_0, y_0, y'_0, \dots, y_0^{(n-1)}$, будет частным решением уравнения (2).

Для выделения его из общего решения надо определить соответствующие значения параметров $C_1, C_2, \dots, C_n$ из соотношений:

$$\begin{aligned} y_0 &= \varphi(x_0, C_1, C_2, \dots, C_n), \\ y'_0 &= \varphi'(x_0, C_1, C_2, \dots, C_n), \\ &\vdots \\ y_0^{(n-1)} &= \varphi^{(n-1)}(x_0, C_1, C_2, \dots, C_n). \end{aligned}$$

Аналогичные рассуждения проводятся для общего интеграла уравнения (2).

Определения общих и частных решений и интегралов уравнений высшего порядка, не разрешенных относительно производных, в данном курсе не рассматриваются.

Основная задача теории дифференциальных уравнений высших порядков — отыскание всех решений каждого данного уравнения и исследование их свойств — является в целом задачей значительно более сложной, чем интегрирование дифференциальных уравнений первого порядка.

§ 11. Уравнения высшего порядка, допускающие понижение порядка

Одним из основных методов, применяемых при интегрировании дифференциальных уравнений высших порядков, является *понижение порядка* уравнения, т. е. сведение уравнения путем замены переменных к другому уравнению, имеющему порядок ниже данного. Понижение порядка уравнения производят даже в тех случаях, когда получаемое при этом уравнение не интегрируется в квадратурах, так как чем ниже порядок уравнения, тем, вообще говоря, легче его решать приближенно численным или графическим методом. Следует заметить, что понижение порядка возможно далеко не для всякого уравнения, в связи с чем представляют интерес некоторые типы неполных дифференциальных уравнений, допускающих понижение порядка.

1. Уравнения вида $y^{(n)} = f(x)$, где $f(x)$ — функция, непрерывная на некотором интервале $a < x < b$ оси OX , не только допускают понижение порядка, но и интегрируются в квадратурах.

Действительно, для любого решения $y = \varphi(x)$ уравнения

$$y^{(n)} = f(x), \quad (1)$$

лежащего в полосе $\{a < x < b, -\infty < y < +\infty\}$, последовательно интегрируя (и тем самым понижая порядок уравнения), получим:

$$\begin{aligned} y^{(n-1)} &= \int f(x) dx + C_1, \\ y^{(n-2)} &= \int (\int f(x) dx + C_1) dx + C_2 = \int dx \int f(x) dx + C_1 x + C_2, \\ &\dots \dots \dots \\ y &= \int dx \int dx \dots \int f(x) dx + C_1 \frac{x^{n-1}}{(n-1)!} + C_2 \frac{x^{n-2}}{(n-2)!} + \dots + C_n, \end{aligned} \quad (2)$$

где каждый из интегралов обозначает одну какую-либо первообразную для подинтегральной функции, а $C_1, C_2, \dots, C_n$ — некоторые постоянные.

Непосредственной подстановкой в уравнение (1) убеждаемся, что функция (2) при любых значениях постоянных $C_1, C_2, \dots, C_n$ удовлетворяет уравнению (1). Отсюда следует, что функция (2) в указанной выше полосе является общим решением данного уравнения.

Пример. Дано уравнение $y''' = e^{2x}$. Требуется найти его частное решение, удовлетворяющее начальным условиям:

$$x_0 = 0, y_0 = 1, y'_0 = -1, y''_0 = 0.$$

Так как функция $f(x) = e^{2x}$ непрерывна всюду, общее решение уравнения в плоскости XOY получится путем трехкратного последовательного интегрирования соотношения $y''' = e^{2x}$:

$$y'' = \frac{1}{2} e^{2x} + C_1, \quad y' = \frac{1}{4} e^{2x} + C_1 x + C_2, \quad y = \frac{1}{8} e^{2x} + C_1 \frac{x^2}{2} + C_2 x + C_3.$$

Общее решение:

$$y = \frac{1}{8} e^{2x} + C_1 \frac{x^2}{2} + C_2 x + C_3,$$

где C_1, C_2, C_3 — произвольные постоянные.

Подставляя в выражения для y, y', y'' начальные значения, получаем для определения постоянных C_1, C_2, C_3 соотношения:

$$1 = \frac{1}{8} + C_3, \quad -1 = \frac{1}{4} + C_2, \quad 0 = -\frac{1}{2} + C_1,$$

откуда находятся $C_1 = -\frac{1}{2}, C_2 = -\frac{5}{4}, C_3 = \frac{7}{8}$.

Искомое частное решение:

$$y = \frac{1}{8} e^{2x} - \frac{1}{4} x^2 - \frac{5}{4} x + \frac{7}{8}.$$

2. Уравнения, не содержащие явно искомой функции и ее производных до порядка $k-1$ включительно, т. е. уравнения вида $F(x, y^{(k)}, y^{(k+1)}, \dots, y^{(n)}) = 0$, допускают понижение порядка на k единиц.

Действительно, введем новую неизвестную функцию, положив $y^{(k)} = z$. Тогда $y^{(k+1)} = z', y^{(k+2)} = z'', \dots, y^{(n)} = z^{(n-k)}$ и уравнение $F(x, y^{(k)}, y^{(k+1)}, \dots, y^{(n)}) = 0$ перепишется в виде:

$$F(x, z, z', \dots, z^{(n-k)}) = 0.$$

Это дифференциальное уравнение порядка $n-k < n$ относительно неизвестной функции z . Таким образом, подстановка $y^{(k)} = z$ понижает порядок данного уравнения на k единиц.

Предположим, что полученное уравнение $(n-k)$ -го порядка проинтегрировано и соотношение

$$z = \psi(x, C_1, C_2, \dots, C_{n-k}),$$

где $C_1, C_2, \dots, C_{n-k}$ — произвольные постоянные, представляет совокупность его решений.

Тогда, подставляя вместо z его значение $y^{(k)}$, мы получим для определения совокупности решений данного уравнения уравнение рассмотренного в п. 1 вида:

$$y^{(k)} = \psi(x, C_1, C_2, \dots, C_{n-k}).$$

Интегрируя его последовательно k раз, получим совокупность решений данного уравнения:

$$y = \varphi(x, C_1, C_2, \dots, C_n).$$

Пример. Дано уравнение $y''' = \frac{y''}{x}$. Определить его решения. Так как уравнение $y''' = \frac{y''}{x}$ не содержит y и y' , то порядок можно понизить, положив $y'' = z$. При этом $y''' = z'$, уравнение запишется в виде $z' = \frac{z}{x}$. Решая его методом разделения переменных, получим:

$$\frac{dz}{z} = \frac{dx}{x}, \quad \ln |z| = \ln |x| + \ln |C_1|, \quad z = C_1 x, \quad C_1 \neq 0.$$

Добавляя решение $z = 0$, утерянное при разделении переменных, найдем совокупность всех решений вспомогательного уравнения в виде:

$$z = C_1 x,$$

где C_1 — произвольная постоянная.

Подставляя y'' вместо z в соотношение $z = C_1 x$, для определения y будем иметь уравнение второго порядка:

$$y'' = C_1 x.$$

Интегрируя его, находим: $y' = C_1 \frac{x^2}{2} + C_2, \quad y = C_1 \frac{x^3}{3!} + C_2 x + C_3$. Это соотношение будет общим интегралом данного уравнения в областях

$$\{x > 0, -\infty < y < +\infty\} \text{ и } \{x < 0, -\infty < y < +\infty\}.$$

3. Уравнения, не содержащие явно независимой переменной, т. е. уравнения вида $F(y, y', \dots, y^{(n)})=0$, допускают понижение порядка на единицу, если за новую независимую переменную взять y , а за новую искомую функцию взять $y' = p$.

Действительно, применяя правило дифференцирования сложной функции, получаем:

$$\begin{aligned} y' = p, \quad y'' = \frac{dy'}{dx} = \frac{dy'}{dy} \cdot \frac{dy}{dx} = \frac{dp}{dy} \cdot p, \quad y''' = \frac{dy''}{dx} = \frac{dy''}{dy} \cdot \frac{dy}{dx} = \\ = \frac{dy''}{dy} \cdot p = \frac{d\left(\frac{dp}{dy} \cdot p\right)}{dy} \cdot p = \frac{d^2p}{dy^2} \cdot p^2 + \left(\frac{dp}{dy}\right)^2 \cdot p \text{ и т. д.} \end{aligned}$$

Каждая из производных от y по x порядка m ($1 \leq m \leq n$) выражается через производные от p по y порядка не выше $m-1$. Подставляя их значения в данное уравнение, получим относительно новой неизвестной функции p дифференциальное уравнение порядка $n-1$:

$$F_1\left(y, p, \frac{dp}{dy}, \dots, \frac{d^{n-1}p}{dy^{n-1}}\right) = 0.$$

Если это уравнение проинтегрировано и $\Phi(y, p, C_1, C_2, \dots, C_{n-1}) = 0$ — совокупность его решений, то для отыскания решений данного дифференциального уравнения остается решить уравнение первого порядка:

$$\Phi(y, y', C_1, C_2, \dots, C_{n-1}) = 0.$$

Пример. Найти решения уравнения $yy'' - y'^2 - 4yy' = 0$. Это уравнение второго порядка, не содержащее явно x . Обозначим через $p = y'$ — новую искомую функцию, а y будем считать новой независимой переменной. Тогда $y'' = \frac{dp}{dy} \cdot p$. В новых переменных данное уравнение имеет вид:

$$py \frac{dp}{dy} - p^2 - 4yp = 0, \text{ т. е. } p \left(y \frac{dp}{dy} - p - 4y \right) = 0.$$

Уравнение $y \frac{dp}{dy} - p - 4y = 0$, т. е. $\frac{dp}{dy} = 4 + \frac{p}{y}$, $y \neq 0$, однородное; полагая $p = uy$, находим:

$$\begin{aligned} u + \frac{du}{dy} \cdot y = 4 + u, \quad du = \frac{4dy}{y}, \quad u = 4 \ln |y| + 4 \ln |C_1|, \\ p = 4y \ln |C_1 y|, \end{aligned}$$

где C_1 — произвольная, отличная от нуля постоянная. Подставив $p = \frac{dy}{dx}$, для определения решений данного уравнения получим соотношение

$$\frac{dy}{dx} = 4y \ln |C_1 y|.$$

Интегрируя его, найдем семейство решений данного уравнения: $\ln |\ln |C_1 y|| = 4x + C_2$, где C_1 и C_2 — произвольные постоянные.

Так как совокупность решений уравнения $py \frac{dp}{dy} - p^2 - 4py = 0$ состоит из решений уравнения $y \frac{dp}{dy} - p - 4y = 0$ и $p = 0$, то к найденному семейству решений данного уравнения следует добавить решения уравнения $y' = 0$, т. е. $y = C$.

Примечание. Если дифференциальное уравнение не содержит явно независимого переменного, искомой функции и ее первых $k-1$ производных, т. е. если это уравнение имеет вид $F(y^{(k)}, y^{(k+1)}, \dots, y^{(n)}) = 0$, то порядок уравнения можно понизить на $k+1$ единицу, применяя последовательно сначала подстановку $y^{(k)} = z$, где z новая неизвестная функция y , а затем выбирая z за новую независимую переменную, а $z' = p$ за новую искомую функцию.

ОБЩАЯ ТЕОРИЯ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ВЫСШЕГО ПОРЯДКА

§ 12. Общие сведения о линейных дифференциальных уравнениях высшего порядка

Дифференциальное уравнение n -го порядка

$$F(x, y, y', \dots, y^{(n)}) = 0$$

называется *линейным*, если функция $F(x, y, y', \dots, y^{(n)})$ линейна относительно искомой функции y и всех ее производных.

Следовательно, любое линейное дифференциальное уравнение n -го порядка может быть записано в виде:

$$A_0(x)y^{(n)} + A_1(x)y^{(n-1)} + \dots + A_n(x)y + A_{n+1}(x) = 0. \quad (1)$$

В случае, если коэффициенты $A_0(x), A_1(x), \dots, A_n(x)$ постоянны, уравнение (1) называется *линейным уравнением с постоянными коэффициентами*. Свободный член $A_{n+1}(x)$ при этом может быть как постоянным, так и зависящим от x .

Линейные дифференциальные уравнения принято записывать в так называемом *приведенном виде*:

$$y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = q(x). \quad (2)$$

Для перехода от уравнения (1) к уравнению (2) достаточно разделить обе части уравнения (1) на $A_n(x)$ и обозначить $p_i(x) = \frac{A_i(x)}{A_0(x)}$, $q(x) = -\frac{A_{n+1}(x)}{A_0(x)}$. Уравнения (1) и (2) эквивалентны там, где $A_0(x) \neq 0$.

Будем в дальнейшем рассматривать уравнение вида (2) с непрерывными на некотором интервале (a, b) оси OX (конечном или бесконечном) коэффициентами $p_1(x), p_2(x), \dots, p_n(x)$ и правой частью $q(x)$.

При таких предположениях уравнение (2) в области $\{a < x < b, -\infty < y < +\infty, -\infty < y' < +\infty, \dots, -\infty < y^{(n-1)} < +\infty\}$ $(n+1)$ -мерного пространства удовлетворяет условиям теоремы существования и единственности решения.

Действительно, так как уравнение (2), разрешенное относительно старшей производной, имеет вид:

$$y^{(n)} = -p_1(x)y^{(n-1)} - \dots - p_n(x)y + q(x),$$

то ¹⁾ функция $f(x, y, y', \dots, y^{(n-1)}) = -p_1(x)y^{(n-1)} - \dots - p_n(x)y + q(x)$ непрерывна в этой области и имеет в ней непрерывные частные производные по $y, y', \dots, y^{(n-1)}$:

$$\frac{\partial f}{\partial y} = -p_n(x), \quad \frac{\partial f}{\partial y'} = -p_{n-1}(x), \quad \dots, \quad \frac{\partial f}{\partial y^{(n-1)}} = -p_1(x).$$

Отсюда следует, что всякая система начальных данных $x_0, y_0, y'_0, \dots, y_0^{(n-1)}$, где $a < x_0 < b$, а $y_0, y'_0, \dots, y_0^{(n-1)}$ — любые числа, определяет в некоторой окрестности точки x_0 единственное решение уравнения (2). Можно доказать, что это решение будет определено не только в окрестности точки x_0 , но и на всем интервале (a, b) .

Уравнение (2) называется *линейным неоднородным уравнением*, или *линейным уравнением с правой частью*, если функция $q(x)$ не равна

¹⁾ См. формулировку теоремы Коши (стр. 448).

тождественно нулю. Если же $q(x) \equiv 0$, то уравнение (2) называется *линейным однородным уравнением*, или *линейным уравнением без правой части*.

Линейные дифференциальные уравнения представляют наиболее простой и изученный тип дифференциальных уравнений высшего порядка. Это объясняется, с одной стороны, тем, что линейные уравнения обладают рядом свойств, значительно облегчающих нахождение и исследование решений, с другой стороны, тем, что решение целого ряда практических задач физики, механики, техники приводит к решению линейных дифференциальных уравнений. В случае, когда та или иная практическая задача приводит к более сложному, нелинейному уравнению, часто применяются приближенные методы решения, позволяющие заменить нелинейное уравнение «ближним» к нему линейным уравнением.

Остановимся подробнее на изучении свойств решений линейных уравнений, а также рассмотрим методы построения их общих решений.

§ 13. Линейные однородные дифференциальные уравнения с произвольными коэффициентами

Рассмотрим линейное однородное уравнение

$$y^{(n)} + p_1(x) y^{(n-1)} + \dots + p_n(x) y = 0 \quad (1)$$

с непрерывными на интервале (a, b) коэффициентами.

По доказанному через каждую точку области σ , представляющей собой полосу $\{a < x < b, -\infty < y < +\infty\}$, проходит решение уравнения (1), и это решение однозначно определяется заданием его начальных условий.

Обозначим через $L(y)$ сокращенно результат применения к функции y совокупности операций, указанных левой частью уравнения (1):

$$L(y) = y^{(n)} + p_1(x) y^{(n-1)} + \dots + p_n(x) y, \quad (2)$$

и будем называть $L(y)$ *линейным дифференциальным оператором*.

Линейный дифференциальный оператор $L(y)$ каждой n раз дифференцируемой функции x ставит в соответствие некоторую функцию x .

Например, если $L(y) = y'' - 5xy' + x^2y$, то

$$L(e^{2x}) = 4e^{2x} - 10xe^{2x} + x^2e^{2x}, \quad L(x^5) = 20x^3 - 25x^5 + x^7 \text{ и т. д.}$$

Линейный дифференциальный оператор (2) обладает следующими свойствами.

1. Если y есть n раз дифференцируемая функция, а C — любое число, то $L(Cy) = CL(y)$.

Действительно,

$$\begin{aligned} L(Cy) &= (Cy)^{(n)} + p_1(x) (Cy)^{(n-1)} + \dots + p_n(x) (Cy) = \\ &= Cy^{(n)} + p_1(x) Cy^{(n-1)} + \dots + p_n(x) Cy = \\ &= C[y^{(n)} + p_1(x) y^{(n-1)} + \dots + p_n(x) y] = CL(y). \end{aligned}$$

2. Если y_1 и y_2 — n раз дифференцируемые функции, то

$$L(y_1 + y_2) = L(y_1) + L(y_2).$$

Действительно,

$$\begin{aligned} L(y_1 + y_2) &= (y_1 + y_2)^{(n)} + p_1(x) (y_1 + y_2)^{(n-1)} + \dots + p_n(x) (y_1 + y_2) = \\ &= [y_1^{(n)} + p_1(x) y_1^{(n-1)} + \dots + p_n(x) y_1] + [y_2^{(n)} + p_1(x) y_2^{(n-1)} + \dots + \\ &\quad + p_n(x) y_2] = L(y_1) + L(y_2). \end{aligned}$$

Примечание. Свойство 2 справедливо для любого конечного числа слагаемых.

Рассмотренное выше линейное однородное уравнение (1) с помощью оператора $L(y)$ может быть сокращенно записано в виде:

$$L(y) = 0. \quad (3)$$

Решения линейного однородного уравнения обладают следующими свойствами:

1. Если функция y_1 является решением уравнения (1), то функция Cy_1 , где C — любое число, также является его решением.

2. Если функции y_1 и y_2 являются решениями уравнения (1), то функция $y_1 + y_2$ также является его решением.

Эти свойства являются непосредственным следствием свойств линейного оператора $L(y)$.

На основании свойств 1 и 2 заключаем, что если $y_1, y_2, \dots, y_n$ — n любых решений уравнения (1), то их линейная комбинация $C_1y_1 + C_2y_2 + \dots + C_ny_n$ с произвольными постоянными коэффициентами $C_1, C_2, \dots, C_n$ также является решением этого уравнения.

Естественно возникает вопрос, не будет ли эта линейная комбинация, зависящая от n произвольных постоянных, являться общим решением уравнения (1) в области σ . Ниже будет показано, что при определенных условиях на решения $y_1, y_2, \dots, y_n$ выражение $C_1y_1 + C_2y_2 + \dots + C_ny_n$ действительно является общим для уравнения (1). Чтобы выяснить, каковы эти условия, введем понятие линейной зависимости и линейной независимости функций.

§ 14. Линейная зависимость и линейная независимость функций. Определитель Вронского и его свойства

Рассмотрим систему из n функций $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$, определенных на одном и том же интервале (a, b) оси OX^1 .

Эти функции называются *линейно зависимыми* на интервале (a, b) , если существуют числа $\alpha_1, \alpha_2, \dots, \alpha_n$, не все равные нулю, и такие, что для всех x интервала (a, b) тождественно выполняется соотношение

$$\alpha_1\varphi_1(x) + \alpha_2\varphi_2(x) + \dots + \alpha_n\varphi_n(x) = 0. \quad (1)$$

Если функции $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ линейно зависимы на интервале (a, b) , то хотя бы одна из этих функций является линейной комбинацией остальных.

Действительно, в соотношении (1) есть коэффициенты, отличные от нуля. Пусть, например, $\alpha_n \neq 0$. Тогда функция $\varphi_n(x)$ есть линейная комбинация остальных функций системы:

$$\varphi_n(x) = \beta_1\varphi_1(x) + \beta_2\varphi_2(x) + \dots + \beta_{n-1}\varphi_{n-1}(x), \text{ где } \beta_i = -\frac{\alpha_i}{\alpha_n}, i = 1, 2, \dots, n-1.$$

Это утверждение можно принять за основное определение линейно зависимых функций. При этом данное выше определение станет следствием²).

Приведем несколько примеров.

Пример 1. Функции $\varphi_1(x) = \sin^2 x$, $\varphi_2(x) = \cos^2 x$, $\varphi_3(x) \equiv 1$ линейно зависимы на любом интервале (a, b) . Действительно, полагая $\alpha_1 = 1$, $\alpha_2 = 1$, $\alpha_3 = -1$, получим на основании известного тригонометрического тождества, что

$$1 \cdot \sin^2 x + 1 \cdot \cos^2 x + (-1) \cdot 1 \equiv 0.$$

¹) Этот интервал может, в частности, быть и бесконечным.

²) Если одна из функций $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$, например $\varphi_n(x)$, является линейной комбинацией остальных, т. е. если

$$\varphi_n(x) = \alpha_1\varphi_1(x) + \alpha_2\varphi_2(x) + \dots + \alpha_{n-1}\varphi_{n-1}(x),$$

то на данном интервале

$$\alpha_1\varphi_1(x) + \alpha_2\varphi_2(x) + \dots + \alpha_{n-1}\varphi_{n-1}(x) + (-1)\varphi_n(x) \equiv 0.$$

Здесь $\alpha_n = -1$, т. е. не равно нулю.

Пример 2. Функции $\varphi_1(x) = \sin^2 x$, $\varphi_2(x) = x$, $\varphi_3(x) = \cos^2 x$, $\varphi_4(x) \equiv 1$, $\varphi_5(x) = e^x$ линейно зависимы на любом интервале (a, b) , так как, полагая $\alpha_1 = 1$, $\alpha_2 = 0$, $\alpha_3 = 1$, $\alpha_4 = -1$, $\alpha_5 = 0$, получим:

$$1 \cdot \sin^2 x + 0 \cdot x + 1 \cdot \cos^2 x + (-1) \cdot 1 + 0 \cdot e^x \equiv 0.$$

Этот пример показывает, что если часть функций системы линейно зависима, то линейно зависимы и все функции.

Пример 3. Функции

$$\varphi_1(x) = \sqrt{x}, \quad \varphi_2(x) = \frac{1}{x}, \quad \varphi_3(x) \equiv 0, \quad \varphi_4(x) = x^2$$

на интервале $0 < x < 1$ линейно зависимы, так как, полагая $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 2$, $\alpha_4 = 0$, получаем:

$$0 \cdot \sqrt{x} + 0 \cdot \frac{1}{x} + 2 \cdot 0 + 0 \cdot x^2 \equiv 0.$$

Примечание. Любая система функций, содержащая функцию, тождественно равную нулю, будет линейно зависимой.

Линейная зависимость двух не равных тождественно нулю на интервале (a, b) функций $\varphi_1(x)$ и $\varphi_2(x)$ равносильна пропорциональности этих функций.

Действительно, если $\alpha_1 \varphi_1(x) + \alpha_2 \varphi_2(x) \equiv 0$, причем $\alpha_1 \neq 0$, то $\varphi_1(x) \equiv -\frac{\alpha_2}{\alpha_1} \varphi_2(x)$.

Функции $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_n(x)$, определенные на интервале (a, b) оси OX , называются *линейно независимыми* на этом интервале, если из соотношения

$$\alpha_1 \varphi_1(x) + \alpha_2 \varphi_2(x) + \dots + \alpha_n \varphi_n(x) \equiv 0,$$

где $\alpha_1, \alpha_2, \dots, \alpha_n$ — числа, следует, что $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

Если функции $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_n(x)$ линейно независимы на интервале (a, b) , то ни одна из них не является линейной комбинацией остальных.

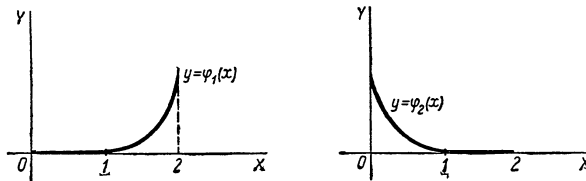


Рис. 344.

Пример. Функции

$$\varphi_0(x) \equiv 1, \quad \varphi_1(x) = x, \quad \varphi_2(x) = x^2, \quad \dots, \quad \varphi_n(x) = x^n,$$

где n — любое натуральное число, линейно независимы на всей числовой оси.

Действительно, если составить линейную комбинацию этих функций с коэффициентами $\alpha_0, \alpha_1, \dots, \alpha_n$, то получится многочлен:

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_n x^n.$$

Многочлен степени не выше n не может иметь более, чем n , действительных корней. Следовательно, тождественное равенство

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_n x^n \equiv 0$$

возможно, только если

$$\alpha_0 = \alpha_1 = \dots = \alpha_n = 0.$$

Рассмотрим еще один пример линейно независимых на интервале $(0, 2)$ функций.

Пусть

$$\varphi_1(x) = \begin{cases} 0, & 0 < x < 1, \\ (x-1)^4, & 1 \leq x < 2, \end{cases} \quad \varphi_2(x) = \begin{cases} (x-1)^4, & 0 < x < 1, \\ 0, & 1 \leq x < 2 \end{cases} \quad (\text{рис. 344}).$$

Подставляя затем $x = \frac{3}{2}$, получим, что

$$\varphi_1\left(\frac{3}{2}\right)=\frac{1}{24}, \quad \varphi_2\left(\frac{3}{2}\right)=0, \quad \alpha_1 \cdot \frac{1}{24}+0 \cdot 0=0, \quad \alpha_1=0.$$

$$\alpha_1 \varphi_1(x) + \alpha_2 \varphi_2(x) \equiv 0$$

Удобный аппарат для исследования некоторых систем функций на линейную зависимость предложен польским математиком Юзефом Вронским.

Определитель n -го порядка

$$W = \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix}$$

Так, например, если $y_1 = \sin x$, $y_2 = e^{-x}$, $y_3 = x^2$, то

$$W(x) = \begin{vmatrix} \sin x & e^{-x} & x^2 \\ \cos x & -e^{-x} & 2x \\ -\sin x & e^{-x} & 2 \end{vmatrix}.$$

Теорема 1. Если функции $y_1, y_2, \dots, y_n$ линейно зависимы на интервале (a, b) , то составленный для них определитель Вронского на этом интервале тождественно равен нулю.

$$y_n = \beta_1 y_1 + \beta_2 y_2 + \dots + \beta_{n-1} y_{n-1},$$

Дифференцируя это тождество последовательно $n-1$ раз, получим:

$$y'_n = \beta_1 y'_1 + \beta_2 y'_2 + \dots + \beta_{n-1} y'_{n-1},$$

$$y_n^{(n-1)} = \beta_1 y_1^{(n-1)} + \beta_2 y_2^{(n-1)} + \dots + \beta_{n-1} y_{n-1}^{(n-1)}.$$

$$W(x) = \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y'_1 & y'_2 & \dots & y'_n \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix} =$$

$$= \begin{vmatrix} y_1 & y_2 & \dots & \beta_1 y_1 & + \beta_2 y_2 & + \dots + \beta_{n-1} y_{n-1} \\ y'_1 & y'_2 & \dots & \beta_1 y'_1 & + \beta_2 y'_2 & + \dots + \beta_{n-1} y'_{n-1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & \beta_1 y_1^{(n-1)} & + \beta_2 y_2^{(n-1)} & + \dots + \beta_{n-1} y_{n-1}^{(n-1)} \end{vmatrix}.$$
$$W(x) = \begin{vmatrix} y_1 & y_2 & \dots & y_{n-1} & 0 \\ y'_1 & y'_2 & \dots & y'_{n-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_{n-1}^{(n-1)} & 0 \end{vmatrix} \equiv 0.$$

Доказательство. Допустим противное. Предположим, что существует точка x_0 , $a < x_0 < b$, в которой определитель Вронского, составленный для функций $y_1, y_2, \dots, y_n$, равен нулю:

$$W(x_0) = \begin{vmatrix} y_1(x_0) & y_2(x_0) & \dots & y_n(x_0) \\ y_1'(x_0) & y_2'(x_0) & \dots & y_n'(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x_0) & y_2^{(n-1)}(x_0) & \dots & y_n^{(n-1)}(x_0) \end{vmatrix} = 0.$$

$$\begin{aligned} \alpha_1 y_1(x_0) + \alpha_2 y_2(x_0) + \dots + \alpha_n y_n(x_0) &= 0, \\ \alpha_1 y_1'(x_0) + \alpha_2 y_2'(x_0) + \dots + \alpha_n y_n'(x_0) &= 0, \\ \vdots & \\ \alpha_1 y_1^{(n-1)}(x_0) + \alpha_2 y_2^{(n-1)}(x_0) + \dots + \alpha_n y_n^{(n-1)}(x_0) &= 0. \end{aligned} \quad (2)$$
$$\tilde{y} = \tilde{a}_1 y_1 + \tilde{a}_2 y_2 + \dots + \tilde{a}_n y_n.$$

⁸⁾ См. теорему о решении линейной однородной системы (гл. I, § 37).

Эта функция, являясь линейной комбинацией решений уравнения $L(y) = 0$, сама будет решением того же уравнения. Так как

$$\tilde{y} = \tilde{\alpha}_1 y_1 + \tilde{\alpha}_2 y_2 + \dots + \tilde{\alpha}_n y_n,$$

$$\tilde{y}' = \tilde{\alpha}_1 y'_1 + \tilde{\alpha}_2 y'_2 + \dots + \tilde{\alpha}_n y'_n,$$

• • • • •

$$\tilde{y}^{(n-1)} = \tilde{a}_1 y_1^{(n-1)} + \tilde{a}_2 y_2^{(n-1)} + \dots + \tilde{a}_n y_n^{(n-1)},$$

то при $x = x_0$ имеем [см. уравнения (2)]:

$$\tilde{y}(x_0) = \tilde{\alpha}_1 y_1(x_0) + \tilde{\alpha}_2 y_2(x_0) + \dots + \tilde{\alpha}_n y_n(x_0) = 0,$$

$$\tilde{y}'(x_0) = \tilde{\alpha}_1 y'_1(x_0) + \tilde{\alpha}_2 y'_2(x_0) + \dots + \tilde{\alpha}_n y'_n(x_0) = 0,$$

.....

$$\tilde{y}^{(n-1)}(x_0) = \tilde{\alpha}_1 y_1^{(n-1)}(x_0) + \tilde{\alpha}_2 y_2^{(n-1)}(x_0) + \dots + \tilde{\alpha}_n y_n^{(n-1)}(x_0) = 0.$$

Это значит, что решение $\tilde{y}$ уравнения $L(y) = 0$ удовлетворяет начальным условиям $x_0, 0, 0, \dots, 0$.

Всякое линейное однородное уравнение имеет так называемое *тривиальное* решение, т. е. решение, тождественно равное нулю: $y=0$ ¹⁾. Это решение также удовлетворяет начальным условиям $x_0, 0, 0, \dots, 0$.

По теореме существования и единственности решения для уравнения $L(y)=0$ задание системы начальных данных однозначно определяет решение, т. е. решение $\tilde{y}$ должно совпадать с решением, тождественно равным нулю:

$$\tilde{y} \equiv 0, \text{ или, что то же, } \tilde{\alpha}_1 y_1 + \tilde{\alpha}_2 y_2 + \dots + \tilde{\alpha}_n y_n \equiv 0.$$

Так как среди чисел $\tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_n$ имеются отличные от нуля, то из последнего соотношения следует, что функции $y_1, y_2, \dots, y_n$ на интервале (a, b) линейно зависимы, что противоречит условию. Таким образом, предположение о том, что определитель $W(x)$ может обращаться в нуль на интервале (a, b) отпадает.

Примечание. Если $n-1$ раз дифференцируемые функции $y_1, y_2, \dots, y_n$ линейно независимы на интервале (a, b) , но не являются на нем решениями линейного однородного дифференциального уравнения, то составленный для них определитель Вронского может быть равен нулю как в отдельных точках интервала (a, b) , так и на всем интервале.

Пример. Как было установлено выше (см. стр. 457), функции

$$\varphi_1(x) = \begin{cases} 0, & 0 < x < 1, \\ (x-1)^4, & 1 \leq x < 2, \end{cases} \quad \varphi_2(x) = \begin{cases} (x-1)^4, & 0 < x < 1, \\ 0, & 1 \leq x < 2, \end{cases}$$

линейно независимы на интервале $(0, 2)$. При этом

$$\varphi_1'(x) = \begin{cases} 0, & 0 < x < 1, \\ 4(x-1)^3, & 1 \leq x < 2, \end{cases} \quad \varphi_2'(x) = \begin{cases} 4(x-1)^3, & 0 < x < 1, \\ 0, & 1 \leq x < 2. \end{cases}$$

Составим для них определитель Вронского:

$$W(x) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{vmatrix}.$$

Для любого x , $0 < x < 1$, имеем:

$$W(x) = \begin{vmatrix} 0 & (x-1)^4 \\ 0 & 4(x-1)^3 \end{vmatrix} = 0.$$

Для любого x , $1 \leq x < 2$, также получаем, что

$$W(x) = \begin{vmatrix} (x-1)^4 & 0 \\ 4(x-1)^3 & 0 \end{vmatrix} = 0,$$

т. е. на интервале $(0, 2)$ $W(x) \equiv 0$.

¹⁾ В этом легко убедиться непосредственной подстановкой функции $y=0$ в уравнение $L(y)=0$.

Из доказанных свойств определителя Вронского следует, что если $y_1, y_2, \dots, y_n$ — определенные на интервале (a, b) решения линейного однородного уравнения $L(y) = 0$, то составленный для них определитель Вронского либо тождественно равен нулю на интервале (a, b) (решения линейно зависимы), либо не равен нулю ни в одной точке интервала (a, b) (решения линейно независимы).

§ 15. Структура общего решения линейного однородного уравнения

Фундаментальной системой решений линейного однородного уравнения n -го порядка на интервале (a, b) называется всякая система из n линейно независимых на этом интервале решений уравнения.

Из теорем предыдущего параграфа следует, что, для того чтобы система решений $y_1, y_2, \dots, y_n$ уравнения

$$y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = 0 \quad (1)$$

была фундаментальной, необходимо и достаточно, чтобы составленный для них определитель Вронского был отличен от нуля хотя бы в одной точке интервала (a, b) .

Легко видеть, что уравнение (1) имеет на интервале (a, b) бесчисленное множество фундаментальных систем решений.

Действительно, рассмотрим любую квадратную таблицу (матрицу) из n^2 чисел

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

с определителем, отличным от нуля¹⁾. Затем закрепим в интервале (a, b) любое значение $x = x_0$ и обозначим через y_1 — решение уравнения (1), удовлетворяющее начальным условиям

$$y_1(x_0) = a_{11}, y_1'(x_0) = a_{21}, \dots, y_1^{(n-1)}(x_0) = a_{n1},$$

через y_2 — решение уравнения (1), удовлетворяющее начальным условиям

$$y_2(x_0) = a_{12}, y_2'(x_0) = a_{22}, \dots, y_2^{(n-1)}(x_0) = a_{n2},$$

и т. д. и, наконец, через y_n — решение уравнения (1), удовлетворяющее начальным условиям

$$y_n(x_0) = a_{1n}, y_n'(x_0) = a_{2n}, \dots, y_n^{(n-1)}(x_0) = a_{nn}.$$

По теореме существования и единственности для уравнения (1) такие решения существуют.

Так как, в силу выбора чисел a_{ik} , $i, k = 1, 2, \dots, n$, определитель

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} y_1(x_0) & y_2(x_0) & \dots & y_n(x_0) \\ y_1'(x_0) & y_2'(x_0) & \dots & y_n'(x_0) \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)}(x_0) & y_2^{(n-1)}(x_0) & \dots & y_n^{(n-1)}(x_0) \end{vmatrix} = W(x_0)$$

отличен от нуля, то решения $y_1, y_2, \dots, y_n$ линейно независимы на интервале (a, b) , т. е. система функций $y_1, y_2, \dots, y_n$ будет фундаментальной.

С помощью фундаментальной системы решений можно решить поставленный выше вопрос о структуре общего интеграла линейного однородного уравнения (1) в области $\sigma \{a < x < b, -\infty < y < +\infty\}$.

¹⁾ Например, удобно взять таблицу $\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}$.

Пример. Линейное однородное уравнение второго порядка $y'' - y = 0$ имеет, как легко видеть, частные решения $y_1 = e^x$ и $y_2 = e^{-x}$. Эти решения на всей числовой оси OX линейно независимы, так как определитель Вронского

$$W(x) = \begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix} = -2,$$

т. е. отличен от нуля.

Общее решение данного уравнения (во всей плоскости XOY) можно, например, записать в виде: $y = C_1 e^x + C_2 e^{-x}$.

§ 16. Нахождение общего решения линейного однородного дифференциального уравнения второго порядка

Как уже говорилось выше, интегрирование линейного однородного уравнения сводится к отысканию фундаментальной системы его решений. Однако даже для уравнения второго порядка с переменными коэффициентами эта задача в общем виде не решается, т. е. не удастся получить ни одной фундаментальной системы решений ни в элементарных функциях, ни в квадратурах. В связи с этим представляет интерес следующая теорема.

Теорема. Если известно одно ненулевое решение $y = y_1$ уравнения

$$y'' + p_1(x)y' + p_2(x)y = 0, \quad (1)$$

то общее решение этого уравнения в области $\{\alpha < x < \beta, -\infty < y < +\infty\}$ может быть найдено с помощью квадратур.

Здесь $\alpha < x < \beta$ — любой интервал, на котором коэффициенты $p_1(x)$ и $p_2(x)$ уравнения непрерывны, а частное решение y_1 не обращается в нуль.

Доказательство. Введем новую неизвестную функцию u , положив $y = uy_1$. Тогда $y' = u'y_1 + uy_1'$, $y'' = u''y_1 + 2u'y_1' + uy_1''$. Подставляя y , y' и y'' в уравнение (1), получим:

$$u''y_1 + 2u'y_1' + uy_1'' + p_1(x)[u'y_1 + uy_1'] + p_2(x)uy_1 = 0,$$

т. е.

$$u''y_1 + u'[2y_1' + p_1(x)y_1] + u[y_1'' + p_1(x)y_1' + p_2(x)y_1] = 0.$$

Так как y_1 — решение уравнения (1), то

$$y_1'' + p_1(x)y_1' + p_2(x)y_1 \equiv 0,$$

а значит, уравнение (1) в новых переменных имеет вид:

$$u''y_1 + u'[2y_1' + p_1(x)y_1] = 0.$$

Порядок этого уравнения понизим, положив $u' = z$, где z — новая искомая функция:

$$z'y_1 + z[2y_1' + p_1(x)y_1] = 0.$$

Решая на интервале $\alpha < x < \beta$ полученное уравнение с разделяющимися переменными, находим совокупность его решений в виде: $z = \frac{C_1}{y_1^2} e^{-\int p_1(x) dx}$, где C_1 — произвольная постоянная. При этом $u' = \frac{C_1}{y_1^2} e^{-\int p_1(x) dx}$, т. е. $u = C_1 \int \frac{1}{y_1^2} e^{-\int p_1(x) dx} dx + C_2$, где C_2 — также произвольная постоянная.

Подставляя найденное значение u в формулу $y = uy_1$, получим совокупность решений уравнения (1):

$$y = C_1 y_1 \int \frac{1}{y_1^2} e^{-\int p_1(x) dx} dx + C_2 y_1. \quad (2)$$

Эта совокупность представляет собой общее решение уравнения (1) в области $\{\alpha < x < \beta, -\infty < y < +\infty\}$, так как решения

$$y_1 \text{ и } y_2 = y_1 \int \frac{1}{y_1^2} e^{-\int p_1(x) dx} dx^1)$$

непропорциональны, а значит, линейно независимы на интервале (α, β) .

Таким образом, если для конкретного линейного однородного уравнения второго порядка удастся подобрать одно ненулевое частное решение, то общее решение уравнения находится указанным в доказательстве теоремы методом.

Пример. Уравнение $y'' + \frac{2}{x}y' + y = 0$ — линейное однородное второго порядка.

Непосредственной подстановкой в уравнение убеждаемся, что функция $y = \frac{\sin x}{x}$ является его решением на любом интервале, не содержащем точки $x = 0$. Применим подстановку $y = \frac{\sin x}{x} u$, где u — новая искомая функция. В новых переменных данное уравнение запишется в виде:

$$\frac{\sin x}{x} u'' + 2 \frac{\cos x}{x} u' = 0$$

или, если положить $u' = z$, в виде:

$$\frac{\sin x}{x} z' + 2 \frac{\cos x}{x} z = 0.$$

Интегрируя полученное уравнение на любом интервале (α, β) , не содержащем точек $x = k\pi$, $k = 0, \pm 1, \pm 2, \dots$, получим:

$$\frac{dz}{z} = -2 \frac{\cos x}{\sin x} dx, \ln |z| = -2 \ln |\sin x| + \ln |C_1|,$$

$z = \frac{C_1}{\sin^2 x}$, где C_1 — произвольная постоянная.

Подставляя $z = u'$ и интегрируя снова, находим:

$u' = \frac{C_1}{\sin^2 x}$, $u = -C_1 \operatorname{ctg} x + C_2$, где C_1 и C_2 — произвольные постоянные.

Функция $y = \frac{\sin x}{x} [-C_1 \operatorname{ctg} x + C_2]$, т. е. $y = -C_1 \frac{\cos x}{x} + C_2 \frac{\sin x}{x}$ будет общим решением данного уравнения в области

$$\{\alpha < x < \beta, -\infty < y < +\infty\}.$$

§ 17. Линейные однородные уравнения высшего порядка с постоянными коэффициентами

Выше уже отмечалось, что фундаментальные системы решений линейных однородных уравнений удастся найти лишь для некоторых простейших типов таких уравнений. Одним из этих типов являются линейные однородные уравнения с постоянными коэффициентами. Покажем, что интегрирование линейных однородных уравнений с постоянными коэффициентами в элементарных функциях сводится к чисто алгебраическим операциям.

Пусть

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y = 0 \quad (1)$$

или, короче,

$$L(y) = 0,$$

где $L(y) = y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y$, — линейное однородное уравнение с действительными коэффициентами.

¹⁾ Эти решения получаются из соотношения (2) соответственно при $C_1 = 0$ и $C_2 = 0$.

Будем искать решения этого уравнения в виде $y = e^{kx}$, где k — некоторое число.

Так как $y' = ke^{kx}$, $y'' = k^2e^{kx}$, ..., $y^{(n)} = k^n e^{kx}$, то

$$L(e^{kx}) = e^{kx} [k^n + a_1 k^{n-1} + \dots + a_n].$$

Многочлен $F(k) = k^n + a_1 k^{n-1} + \dots + a_n$ называется *характеристическим многочленом* дифференциального уравнения (1). Формально этот многочлен получается из левой части уравнения (1), если все производные искомой функции заменить соответствующими степенями k .

Для того чтобы функция $y = e^{kx}$ являлась решением уравнения (1), необходимо и достаточно, чтобы

$$L(e^{kx}) = 0, \text{ т. е. } e^{kx} F(k) = 0.$$

Множитель e^{kx} отличен от нуля, следовательно, число k должно удовлетворять уравнению

$$F(k) = 0. \quad (2)$$

Уравнение $F(k) = 0$ называется *характеристическим уравнением* (с неизвестной k), соответствующим данному дифференциальному уравнению (1).

Таким образом, для того чтобы функция $y = e^{kx}$ была решением дифференциального уравнения (1), необходимо и достаточно, чтобы число k было корнем соответствующего характеристического уравнения (2).

Характеристическое уравнение (2)

$$k^n + a_1 k^{n-1} + \dots + a_n = 0$$

есть алгебраическое уравнение n -й степени относительно k , следовательно, по основной теореме алгебры оно имеет n корней: $k_1, k_2, \dots, k_n$. Каждому из этих корней соответствует решение дифференциального уравнения (1).

При изучении вопроса об общем решении уравнения (1) (во всей плоскости XOY) приходится рассматривать следующие возможные случаи.

1. Все корни характеристического уравнения (2) действительны и различны.

Функции $y_1 = e^{k_1 x}$, $y_2 = e^{k_2 x}$, ..., $y_n = e^{k_n x}$ являются по доказанному решениями дифференциального уравнения (1). Эти решения составляют фундаментальную систему решений (на всей числовой оси).

Для доказательства составим определитель Вронского:

$$W(x) = \begin{vmatrix} e^{k_1 x} & e^{k_2 x} & \dots & e^{k_n x} \\ k_1 e^{k_1 x} & k_2 e^{k_2 x} & \dots & k_n e^{k_n x} \\ \dots & \dots & \dots & \dots \\ k_1^{n-1} e^{k_1 x} & k_2^{n-1} e^{k_2 x} & \dots & k_n^{n-1} e^{k_n x} \end{vmatrix} = e^{(k_1 + k_2 + \dots + k_n)x} V,$$

где

$$V = \begin{vmatrix} 1 & 1 & \dots & 1 \\ k_1 & k_2 & \dots & k_n \\ \dots & \dots & \dots & \dots \\ k_1^{n-1} & k_2^{n-1} & \dots & k_n^{n-1} \end{vmatrix}$$

— определитель, известный под названием *определителя Вандермонда*. Он равен произведению всевозможных разностей первых степеней входящих в него элементов (см. гл. I, § 36):

$$V = (k_2 - k_1)(k_3 - k_1) \dots (k_n - k_1)(k_3 - k_2) \dots (k_n - k_2) \dots (k_n - k_{n-1}).$$

Так как корни $k_1, k_2, \dots, k_n$ различны, то определитель Вандермонда, а значит, и определитель Вронского, отличны от нуля.

Таким образом, в случае различных действительных корней $k_1, k_2, \dots, k_n$ характеристического уравнения (2) решения

$$y_1 = e^{k_1 x}, y_2 = e^{k_2 x}, \dots, y_n = e^{k_n x}$$

составляют фундаментальную систему, а значит, общее решение уравнения (2) можно записать в виде:

$$y = C_1 e^{k_1 x} + C_2 e^{k_2 x} + \dots + C_n e^{k_n x},$$

где $C_1, C_2, \dots, C_n$ — произвольные постоянные.

Пример 1. Дано уравнение $y'' - 3y' + 2y = 0$. Характеристическое уравнение имеет вид $k^2 - 3k + 2 = 0$, его корни $k_1 = 1, k_2 = 2$ — действительные и различные, соответствующие частные решения уравнения: $y_1 = e^x, y_2 = e^{2x}$. Общее решение: $y = C_1 e^x + C_2 e^{2x}$, где C_1 и C_2 — произвольные постоянные.

Пример 2. Дано уравнение $y''' - 7y'' + 6y' = 0$. Составляем характеристическое уравнение: $k^3 - 7k^2 + 6k = 0$, находим его корни $k_1 = 0, k_2 = 1, k_3 = 6$. Корни действительные и различные. Им соответствуют решения: $y_1 = 1, y_2 = e^x, y_3 = e^{6x}$. Общее решение уравнения: $y = C_1 + C_2 e^x + C_3 e^{6x}$, где C_1, C_2, C_3 — произвольные постоянные.

2. Все корни характеристического уравнения действительны, но среди них встречаются кратные корни.

В этом случае среди чисел $k_1, k_2, \dots, k_n$ различных корней будет меньше, чем n , и соответственно среди решений $e^{k_1 x}, e^{k_2 x}, \dots, e^{k_n x}$ различных решений также будет меньше, чем n . Это значит, что система из этих решений уже не будет фундаментальной.

Для получения недостающих решений можно использовать свойства оператора:

$$L(y) = y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y$$

с постоянными коэффициентами.

Равенство $L(e^{kx}) = e^{kx} F(k)$ имеет место для всех значений числа k . Продифференцируем его m раз по k :

$$\frac{\partial^m L(e^{kx})}{\partial k^m} = \frac{\partial^m [e^{kx} F(k)]}{\partial k^m}.$$

При этом

$$\frac{\partial^m L(e^{kx})}{\partial k^m} = \frac{\partial}{\partial k^m} [(e^{kx})^n + a_1 (e^{kx})^{(n-1)} + \dots + a_n e^{kx}].$$

Учитывая свойства производных, а также теорему о независимости смешанной производной от последовательности дифференцирований, получаем:

$$\begin{aligned} \frac{\partial^m L(e^{kx})}{\partial k^m} &= \left(\frac{\partial^m e^{kx}}{\partial k^m} \right)^{(n)} + a_1 \left(\frac{\partial^m e^{kx}}{\partial k^m} \right)^{(n-1)} + \dots + a_n \frac{\partial^m e^{kx}}{\partial k^m} = \\ &= L \left(\frac{\partial^m e^{kx}}{\partial k^m} \right) = L(x^m e^{kx}). \end{aligned}$$

С другой стороны, применяя правило Лейбница для вычисления производной порядка m от произведения двух функций, находим:

$$\begin{aligned} \frac{\partial^m [e^{kx} F(k)]}{\partial k^m} &= \frac{\partial^m e^{kx}}{\partial k^m} F(k) + m \frac{\partial^{m-1} e^{kx}}{\partial k^{m-1}} F'(k) + \\ &+ C_m^2 \frac{\partial^{m-2} e^{kx}}{\partial k^{m-2}} F''(k) + \dots + e^{kx} F^{(m)}(k) = \\ &= e^{kx} [x^m F(k) + mx^{m-1} F'(k) + C_m^2 x^{m-2} F''(k) + \dots + F^{(m)}(k)]. \end{aligned}$$

Приравнявая полученные результаты, приходим к формуле

$$L(x^m e^{kx}) = e^{kx} [x^m F(k) + mx^{m-1} F'(k) + C_m^2 x^{m-2} F''(k) + \dots + F^{(m)}(k)], \quad (3)$$

справедливой для любого натурального числа m и любого числа k .

Вернемся к поставленной выше задаче. Пусть k_1 — корень уравнения (2) кратности m_1 . Это значит, что характеристический многочлен $F(k)$ можно записать в виде:

$$F(k) = (k - k_1)^{m_1} \Phi(k), \text{ где } \Phi(k_1) \neq 0.$$

Так как всякий кратный корень многочлена является корнем кратности на единицу меньшей для его производной¹⁾, заключаем, что

$$F(k_1) = F'(k_1) = \dots = F^{(m_1-1)}(k_1) = 0, \text{ а } F^{(m_1)}(k_1) \neq 0.$$

Подставляя в формулу (3) вместо k число k_1 , а вместо m последовательно числа $0, 1, 2, \dots, m_1 - 1$, получим, что $L(x^{m_1} e^{k_1 x}) = 0$. Это значит, что функции

$$e^{k_1 x}, \quad x e^{k_1 x}, \quad x^2 e^{k_1 x}, \dots, \quad x^{m_1-1} e^{k_1 x}$$

являются решениями данного дифференциального уравнения. Таким образом, корню k_1 характеристического уравнения (2) кратности m_1 ставится в соответствие ровно m_1 различных решений дифференциального уравнения.

Рассуждая указанным образом для каждого простого и кратного корня характеристического уравнения, составим систему из n различных решений дифференциального уравнения (1). Можно показать, что эта система решений будет фундаментальной на интервале $-\infty < x < +\infty$. Линейная комбинация этих решений с произвольными коэффициентами дает общее решение уравнения (1) во всей плоскости.

Пример 3. Дано уравнение $y'' - 4y' + 4y = 0$. Его характеристическое уравнение $k^2 - 4k + 4 = 0$ имеет кратные корни $k_1 = k_2 = 2$. Соответствующие им частные решения уравнений имеют вид: $y_1 = e^{2x}$, $y_2 = x e^{2x}$. Общее решение: $y = C_1 e^{2x} + C_2 x e^{2x}$, где C_1 и C_2 — произвольные постоянные.

Найдем частное решение этого уравнения с начальными данными $0, -1, 1$.

Так как $y = C_1 e^{2x} + C_2 x e^{2x}$, то $y' = 2C_1 e^{2x} + C_2 e^{2x}(1 + 2x)$.

При $x = 0$ имеем $-1 = C_1$, $1 = 2C_1 + C_2$, откуда $C_1 = -1$, $C_2 = 3$.

Искомое частное решение имеет вид: $y = e^{-2x} + 3x e^{2x}$.

Пример 4. Дано уравнение $y^{(5)} - 9y''' = 0$. Характеристическое уравнение имеет вид: $k^5 - 9k^3 = 0$, его корни $k_1 = k_2 = k_3 = 0$, $k_4 = 3$, $k_5 = -3$. Корень $k_1 = 0$ — трехкратный, корни $k_4 = 3$, $k_5 = -3$ — простые. Соответствующие частные решения уравнения: $y_1 = 1$, $y_2 = x$, $y_3 = x^2$, $y_4 = e^{3x}$, $y_5 = e^{-3x}$. Общее решение: $y = C_1 + C_2 x + C_3 x^2 + C_4 e^{3x} + C_5 e^{-3x}$, где C_1, C_2, C_3, C_4 и C_5 — произвольные постоянные.

3. Среди корней характеристического уравнения (2) встречаются мнимые корни.

В этом случае для решения вопроса о фундаментальной системе решений уравнения (1) требуется ряд элементарных сведений из теории комплексных функций действительного аргумента. Приведем необходимые сведения (подробнее эти вопросы изучаются в гл. X).

Если каждому значению действительной переменной x (из некоторого множества) поставлено в соответствие комплексное число $y = u + iv$, то говорят, что задана *комплексная функция действительного аргумента* x и пишут:

$$y = f(x) \text{ или } y = u(x) + iv(x).$$

При этом функции $u(x)$ и $v(x)$ называются соответственно *действительной* и *мнимой частью* функции $y = f(x)$.

Для комплексных функций действительной переменной x можно ввести понятие предела, непрерывности, производной аналогично тому, как это делалось в действительном случае. Из определения производной, в частности, следует, что

$$y' = u'(x) + iv'(x), \dots, y^{(n)} = u^{(n)}(x) + iv^{(n)}(x).$$

¹⁾ Действительно, если $F(k) = (k - k_1)^{m_1} \Phi(k)$, где $\Phi(k_1) \neq 0$, то $F'(k) = (k - k_1)^{m_1-1} [m_1 \Phi(k) + (k - k_1) \Phi'(k)] = (k - k_1)^{m_1-1} \Psi(k)$, где $\Psi(k_1) = m_1 \Phi(k_1) \neq 0$. Это и значит, что $k = k_1$ — корень многочлена $F'(k)$ кратности $m_1 - 1$.

Рассмотрим линейное однородное уравнение (1) с действительными коэффициентами:

$$L(y) = y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y.$$

Комплексная функция действительной переменной $y = u(x) + iv(x)$ называется решением этого уравнения, если

$$L(y) \equiv 0, \text{ т. е. } L(u(x) + iv(x)) \equiv 0.$$

Подставляя $y = u(x) + iv(x)$ в уравнение (1), получим:

$$\begin{aligned} (u + iv)^{(n)} + a_1 (u + iv)^{(n-1)} + \dots + a_n (u + iv) &= \\ = [u^{(n)} + a_1 u^{(n-1)} + \dots + a_n u] + i [v^{(n)} + a_1 v^{(n-1)} + \dots + a_n v] &= \\ = L(u) + iL(v) \equiv 0. \end{aligned}$$

Так как комплексное число равно нулю только тогда, когда его действительная и мнимая части равны нулю, заключаем, что

$$L(u(x)) \equiv 0 \text{ и } L(v(x)) \equiv 0.$$

Таким образом, если функция $y = u(x) + iv(x)$ есть решение уравнения (1), то ее действительная и мнимая части также являются решениями уравнения (1).

Определим для любых действительных чисел α и β комплексную показательную функцию действительного аргумента x равенством:

$$e^{(\alpha + i\beta)x} = e^{\alpha x} (\cos \beta x + i \sin \beta x).$$

В частности, если $\alpha = 0$, отсюда следует, что для любого действительного числа β имеет место равенство

$$e^{i\beta x} = \cos \beta x + i \sin \beta x.$$

Подставляя вместо числа β число $-\beta$, получим равенство

$$e^{-i\beta x} = \cos \beta x - i \sin \beta x.$$

Из этих равенств как следствие получаем, что

$$\cos \beta x = \frac{e^{i\beta x} + e^{-i\beta x}}{2}, \quad \sin \beta x = \frac{e^{i\beta x} - e^{-i\beta x}}{2i}.$$

Функция $e^{(\alpha + i\beta)x}$ для любого значения x имеет производные всех порядков, причем:

$$\begin{aligned} (e^{(\alpha + i\beta)x})' &= (\alpha + i\beta) e^{(\alpha + i\beta)x}, \\ (e^{(\alpha + i\beta)x})'' &= (\alpha + i\beta)^2 e^{(\alpha + i\beta)x}, \\ \vdots &\vdots \\ (e^{(\alpha + i\beta)x})^{(n)} &= (\alpha + i\beta)^n e^{(\alpha + i\beta)x}. \end{aligned}$$

Вернемся к решению линейного однородного уравнения (1) в случае, когда среди корней его характеристического уравнения встречаются мнимые корни.

Пусть $k_1 = \alpha_1 + i\beta_1$ — простой корень характеристического уравнения (2). Так как это уравнение имеет действительные коэффициенты, то среди его корней найдется корень, комплексно сопряженный корню k_1 . Обозначим его через k_2 :

$$k_2 = \alpha_1 - i\beta_1.$$

Корням k_1 и k_2 соответствуют комплексные решения

$$y_1 = e^{(\alpha_1 + i\beta_1)x} \text{ и } y_2 = e^{(\alpha_1 - i\beta_1)x}$$

дифференциального уравнения (1).

По доказанному действительные и мнимые части этих решений, в свою очередь, являются решениями уравнения (1). Так как

$$y_1 = e^{\alpha_1 x} \cos \beta_1 x + i e^{\alpha_1 x} \sin \beta_1 x, \text{ а } y_2 = e^{\alpha_1 x} \cos \beta_1 x - i e^{\alpha_1 x} \sin \beta_1 x,$$

то это значит, что функции

$$e^{\alpha_1 x} \cos \beta_1 x, e^{\alpha_1 x} \sin \beta_1 x \text{ и } -e^{\alpha_1 x} \sin \beta_1 x$$

будут действительными решениями уравнения (1). Отбрасывая последнюю из них, получаем два действительных решения уравнения (1):

$$\tilde{y}_1 = e^{\alpha_1 x} \cos \beta_1 x \text{ и } \tilde{y}_2 = e^{\alpha_1 x} \sin \beta_1 x,$$

соответствующие двум простым комплексно сопряженным корням $k_{1,2} = \alpha_1 \pm i\beta_1$ характеристического уравнения ¹⁾.

Если число $k_1 = \alpha_1 + i\beta_1$ — корень характеристического уравнения (2) кратности m_1 , то комплексно сопряженное число $k_2 = \alpha_1 - i\beta_1$ также является корнем уравнения (2) кратности m_1 . Этим корням соответствует $2m_1$ комплексных решений дифференциального уравнения (1):

$$e^{(\alpha_1 + i\beta_1)x}, xe^{(\alpha_1 + i\beta_1)x}, \dots, x^{m_1-1} e^{(\alpha_1 + i\beta_1)x}, \\ e^{(\alpha_1 - i\beta_1)x}, xe^{(\alpha_1 - i\beta_1)x}, \dots, x^{m_1-1} e^{(\alpha_1 - i\beta_1)x}.$$

По комплексным решениям, отделяя их действительные и мнимые части, можно составить систему из $2m_1$ действительных решений того же уравнения (1):

$$e^{\alpha_1 x} \cos \beta_1 x, xe^{\alpha_1 x} \cos \beta_1 x, \dots, x^{m_1-1} e^{\alpha_1 x} \cos \beta_1 x, \\ e^{\alpha_1 x} \sin \beta_1 x, xe^{\alpha_1 x} \sin \beta_1 x, \dots, x^{m_1-1} e^{\alpha_1 x} \sin \beta_1 x.$$

Аналогичным образом поступаем с любой парой комплексно сопряженных корней характеристического уравнения.

Рассмотренные выше случаи позволяют сформулировать следующее общее правило решения линейного однородного дифференциального уравнения с постоянными коэффициентами.

1. Составляем характеристическое уравнение и отыскиваем все его корни.

2. Находим частные решения данного дифференциального уравнения; при этом:

а) каждому простому действительному корню k характеристического уравнения ставится в соответствие решение e^{kx} ,

б) каждому m -кратному действительному корню k характеристического уравнения ставится в соответствие m решений:

$$e^{kx}, xe^{kx}, \dots, x^{m-1} e^{kx},$$

в) каждой паре простых комплексно сопряженных корней $\alpha \pm i\beta$ характеристического уравнения ставятся в соответствие два решения:

$$e^{\alpha x} \cos \beta x \text{ и } e^{\alpha x} \sin \beta x,$$

г) каждой паре m -кратных комплексно сопряженных корней $\alpha \pm i\beta$ характеристического уравнения ставится в соответствие $2m$ решений:

$$e^{\alpha x} \cos \beta x, xe^{\alpha x} \cos \beta x, \dots, x^{m-1} e^{\alpha x} \cos \beta x, \\ e^{\alpha x} \sin \beta x, xe^{\alpha x} \sin \beta x, \dots, x^{m-1} e^{\alpha x} \sin \beta x.$$

Можно показать, что полученная таким образом совокупность решений образует фундаментальную систему решений уравнения на участке $-\infty < x < +\infty$.

3. Составляем линейную комбинацию найденных решений.

Эта линейная комбинация решений с произвольными коэффициентами даст общее решение уравнения в плоскости XOY .

¹⁾ Отброшенное решение — $e^{\alpha_1 x} \sin \beta_1 x$ пропорционально решению $\tilde{y}_2$.

Пример 5. Дано уравнение $y''' - y = 0$. Характеристическое уравнение $k^3 - 1 = 0$ имеет корни $k_1 = 1$, $k_{2,3} = -\frac{1}{2} \pm i \frac{\sqrt{3}}{2}$. Этим корням соответствуют решения

$$y_1 = e^x, \quad y_2 = e^{-\frac{x}{2}} \cos \frac{\sqrt{3}}{2} x, \quad y_3 = e^{-\frac{x}{2}} \sin \frac{\sqrt{3}}{2} x.$$

Общее решение уравнения:

$$y = C_1 e^x + e^{-\frac{x}{2}} \left[C_2 \cos \frac{\sqrt{3}}{2} x + C_3 \sin \frac{\sqrt{3}}{2} x \right],$$

где C_1, C_2, C_3 — произвольные постоянные.

Пример 6. Если уравнение $L(y) = 0$ с постоянными коэффициентами имеет корнями характеристического уравнения числа $k_{1,2,3,4} = -1$, $k_{5,6} = 2 + 3i$, $k_{7,8} = 2 - 3i$, то общее решение этого уравнения имеет вид:

$$y = C_1 e^{-x} + C_2 x e^{-x} + C_3 x^2 e^{-x} + C_4 x^3 e^{-x} + C_5 e^{2x} \cos 3x + \\ + C_6 x e^{2x} \cos 3x + C_7 e^{2x} \sin 3x + C_8 x e^{2x} \sin 3x,$$

где $C_1, C_2, \dots, C_8$ — произвольные постоянные.

§ 18. Линейные неоднородные дифференциальные уравнения с произвольными коэффициентами. Структура общего решения

Рассмотрим линейное неоднородное уравнение n -го порядка:

$$y^{(n)} + p_1(x) y^{(n-1)} + \dots + p_n(x) y = f(x), \quad (1)$$

коэффициенты и правая часть которого непрерывны на некотором (конечном или бесконечном) интервале (a, b) .

Обозначив через $L(y)$ оператор

$$L(y) = y^{(n)} + p_1(x) y^{(n-1)} + \dots + p_n(x) y,$$

можно записать уравнение (1) сокращенно в виде:

$$L(y) = f(x).$$

Выясним структуру общего решения этого уравнения в области σ :

$$\{a < x < b, -\infty < y < +\infty\}.$$

Теорема. Общее решение линейного неоднородного уравнения (1) в области σ есть сумма любого его решения и общего решения соответствующего однородного уравнения в той же области.

Доказательство. Пусть Y — какое-нибудь определенное решение уравнения (1) в области σ : $L(Y) \equiv f(x)$.

Обозначим через $y_1, y_2, \dots, y_n$ фундаментальную систему решений линейного однородного уравнения

$$L(y) = 0, \quad (2)$$

соответствующего уравнению (1). Тогда общее решение уравнения $L(y) = 0$ в области σ запишется в виде:

$$y = C_1 y_1 + C_2 y_2 + \dots + C_n y_n,$$

где $C_1, C_2, \dots, C_n$ — произвольные постоянные.

Покажем, что выражение

$$Y + C_1 y_1 + C_2 y_2 + \dots + C_n y_n, \quad (3)$$

зависящее от n произвольных постоянных, является общим решением уравнения (1) в области σ .

Действительно, при любых значениях постоянных $C_1, C_2, \dots, C_n$ выражение (3) является решением данного уравнения, так как

$$L(Y + C_1 y_1 + C_2 y_2 + \dots + C_n y_n) = L(Y) + C_1 L(y_1) + C_2 L(y_2) + \dots + \\ + C_n L(y_n) = L(Y) \equiv f(x).$$

Остается доказать, что любое решение $\tilde{y}$ уравнения (1), лежащее в области σ , можно записать в виде (3) при некоторых значениях $C_1, C_2, \dots, C_n$.

Так как $\tilde{y}$ — решение уравнения (1), то $L(\tilde{y}) \equiv f(x)$. При этом разность решений $\tilde{y}$ и Y уравнения (1) оказывается решением линейного однородного уравнения (2):

$$L(\tilde{y} - Y) = L(\tilde{y}) - L(Y) \equiv f(x) - f(x) = 0,$$

а значит, она может быть записана в виде: $C_1 y_1 + C_2 y_2 + \dots + C_n y_n$ при некоторых конкретных значениях $C_1, C_2, \dots, C_n$:

$$\tilde{y} - Y = \tilde{C}_1 y_1 + \tilde{C}_2 y_2 + \dots + \tilde{C}_n y_n.$$

Отсюда следует, что

$$\tilde{y} = Y + \tilde{C}_1 y_1 + \tilde{C}_2 y_2 + \dots + \tilde{C}_n y_n.$$

Примечание. Так как любое решение уравнения (1) в области σ (включая и Y) можно получить из общего решения (3) при специальных значениях постоянных, то каждое такое решение уравнения (1) будет для него частным решением. Поэтому доказанную теорему обычно формулируют так: общее решение линейного неоднородного уравнения в области σ есть сумма любого его частного решения и общего решения соответствующего однородного уравнения в этой области.

Пример. Уравнение $y''' - 4y' = x$ — линейное неоднородное третьего порядка. Общее решение соответствующего однородного уравнения имеет вид: $y = C_1 + C_2 e^{2x} + C_3 e^{-2x}$. Частное решение данного уравнения: $y = -\frac{x^2}{8}$ ¹⁾. Общее решение данного уравнения (в плоскости XOY):

$$y = -\frac{x^2}{8} + C_1 + C_2 e^{2x} + C_3 e^{-2x},$$

где C_1, C_2 и C_3 — произвольные постоянные.

При решении линейных неоднородных уравнений с комбинированными правыми частями иногда используется следующая **теорема**: если y_1 и y_2 — частные решения уравнений $L(y) = f_1(x)$ и $L(y) = f_2(x)$ в области σ , то $y_1 + y_2$ есть частное решение уравнения $L(y) = f_1(x) + f_2(x)$ в той же области.

Доказательство вытекает из соотношений: $L(y_1) \equiv f_1(x)$, $L(y_2) \equiv f_2(x)$ и свойства оператора $L(y)$:

$$L(y_1 + y_2) = L(y_1) + L(y_2) \equiv f_1(x) + f_2(x).$$

Пример. Дано уравнение $y'' + y = x - \sin 2x$. Рассмотрим два вспомогательных уравнения: $y'' + y = x$ и $y'' + y = -\sin 2x$.

Частные решения этих уравнений соответственно равны

$$y_1 = x, \quad y_2 = \frac{1}{3} \sin 2x.$$

Общее решение однородного уравнения $y'' + y = 0$ имеет вид: $y = C_1 \sin x + C_2 \cos x$, где C_1 и C_2 — произвольные постоянные.

На основании доказанных теорем, общее решение данного уравнения (в плоскости XOY) можно записать в виде:

$$y = x + \frac{1}{3} \sin 2x + C_1 \sin x + C_2 \cos x.$$

¹⁾ Это частное решение находится методом подбора. Подробнее о методе подбора см. § 19.

Ниже мы остановимся подробно на некоторых правилах подбора частных решений линейных неоднородных уравнений с постоянными коэффициентами. В общем случае задача подбора частного решения представляет значительные трудности, в связи с чем становится интересным изучение так называемого *метода вариации постоянных*, позволяющего найти общее решение линейного неоднородного уравнения в квадратурах по известной фундаментальной системе решений соответствующего однородного уравнения.

Теорема. Общее решение линейного неоднородного уравнения (1):

$$L(y) = y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = f(x)$$

в области σ может быть найдено в квадратурах, если известно общее решение соответствующего однородного уравнения в той же области.

Доказательство. Пусть $y_1, y_2, \dots, y_n$ — фундаментальная система решений уравнения (2):

$$L(y) = 0$$

на интервале $a < x < b$. Общее решение такого уравнения в области σ запишется в виде:

$$y = C_1 y_1 + C_2 y_2 + \dots + C_n y_n$$

или, короче, в виде:

$$y = \sum_{i=1}^n C_i y_i,$$

где $C_1, C_2, \dots, C_n$ — произвольные постоянные.

Будем искать решение линейного неоднородного уравнения (1) в той же форме. При этом $C_1, C_2, \dots, C_n$ придется считать функциями от x :

$$y = \sum_{i=1}^n C_i(x) y_i. \quad (3')$$

Подберем эти n неизвестных функций $C_1(x), C_2(x), \dots, C_n(x)$ таким образом, чтобы выражение (3') являлось решением уравнения (1). Для этого необходимо, чтобы на интервале (a, b) тождественно выполнялось условие

$$L\left(\sum_{i=1}^n C_i(x) y_i\right) = f(x). \quad (4)$$

Остальные $(n-1)$ условий, необходимые для определения n неизвестных функций, можно задавать произвольно, лишь бы полученная система оказалась совместной. Эти дополнительные условия будем подбирать постепенно таким образом, чтобы условие (4), получающееся в результате

подстановки функции $y = \sum_{i=1}^n C_i(x) y_i$ и всех ее производных до n -го порядка включительно в уравнение (1), имело наиболее простой вид.

Вычислим эти производные.

Если $y = \sum_{i=1}^n C_i(x) y_i$, то $y' = \sum_{i=1}^n C_i(x) y_i' + \sum_{i=1}^n C_i'(x) y_i$. В качестве

первого из дополнительных условий возьмем уравнение $\sum_{i=1}^n C_i'(x) y_i = 0$.

При этом для y' получим выражение

$$y' = \sum_{i=1}^n C_i(x) y_i', \quad (5)$$

которое имеет такой же вид, как и в случае постоянных C_i .

Для нахождения y'' дифференцируем соотношение (5) по x :

$$y'' = \sum_{i=1}^n C_i(x) y_i'' + \sum_{i=1}^n C_i'(x) y_i'.$$

В качестве второго дополнительного условия берем уравнение

$$\sum_{i=1}^n C_i'(x) y_i' = 0.$$

При этом

$$y'' = \sum_{i=1}^n C_i(x) y_i'' \text{ и т. д.}$$

Последнее $(n-1)$ -е дополнительное условие налагается при $(n-1)$ -м шаге:

$$y^{(n-1)} = \sum_{i=1}^n C_i(x) y_i^{(n-1)} + \sum_{i=1}^n C_i'(x) y_i^{(n-2)}.$$

Оно имеет вид:

$$\sum_{i=1}^n C_i'(x) y_i^{(n-2)} = 0.$$

При этом условии

$$y^{(n-1)} = \sum_{i=1}^n C_i(x) y_i^{(n-1)}. \quad (6)$$

Дифференцируя соотношение (6) по x , находим:

$$y^{(n)} = \sum_{i=1}^n C_i(x) y_i^{(n)} + \sum_{i=1}^n C_i'(x) y_i^{(n-1)}.$$

Подставив выражение $y = \sum_{i=1}^n C_i(x) y_i$ и все его найденные производные в левую часть уравнения (1), получим:

$$\begin{aligned} L\left(\sum_{i=1}^n C_i(x) y_i\right) &= \sum_{i=1}^n C_i(x) y_i^{(n)} + \sum_{i=1}^n C_i'(x) y_i^{(n-1)} + \\ &+ p_1(x) \sum_{i=1}^n C_i(x) y_i^{(n-1)} + \dots + p_n(x) \sum_{i=1}^n C_i(x) y_i = \\ &= \sum_{i=1}^n C_i'(x) y_i^{(n-1)} + \sum_{i=1}^n C_i(x) [y_i^{(n)} + p_1(x) y_i^{(n-1)} + \dots + p_n(x) y_i] = \\ &= \sum_{i=1}^n C_i'(x) y_i^{(n-1)} + \sum_{i=1}^n C_i(x) L(y_i) = \sum_{i=1}^n C_i'(x) y_i^{(n-1)}, \end{aligned}$$

так как

$$L(y_1) = L(y_2) = \dots = L(y_n) = 0.$$

Условие (4) [n -е условие для определения неизвестных функций $C_1(x)$, $C_2(x)$, ..., $C_n(x)$] при этом запишется в виде:

$$\sum_{i=1}^n C_i'(x) y_i^{(n-1)} = f(x).$$

Таким образом, для того чтобы выражение

$$y = \sum_{i=1}^n C_i(x) y_i$$

1) Найти фундаментальную систему решений соответствующего однородного уравнения (2) на том интервале, где непрерывны коэффициенты и правая часть уравнения (1).

2) Установить вид общего решения уравнения (2) в области σ .

3) Записать общее решение уравнения (1) в той же форме (3'), но с переменными коэффициентами.

4) Составить для определения производных от этих переменных коэффициентов систему уравнений (7).

5) Решив систему и проинтегрировав найденные производные, найти переменные коэффициенты.

5) Подставить найденные коэффициенты в формулу (3').

Пример. Будем искать общее решение уравнения $y''' + y' = \frac{1}{\sin x}$ в области $\sigma \{a < x < b, -\infty < y < +\infty\}$, где (a, b) — любой интервал, не содержащий точек $k\pi$, $k = 0, \pm 1, \pm 2, \dots$. Соответствующее однородное уравнение $y''' + y' = 0$ имеет фундаментальную систему решений: $y_1 = 1$, $y_2 = \cos x$, $y_3 = \sin x$. Общее решение однородного уравнения:

$$y = C_1 + C_2 \cos x + C_3 \sin x,$$

где C_1, C_2, C_3 — произвольные постоянные.

Ищем общее решение данного уравнения в виде:

$$y = C_1(x) + C_2(x) \cos x + C_3(x) \sin x.$$

Для определения $C_1'(x)$, $C_2'(x)$, $C_3'(x)$ имеем систему уравнений:

$$\begin{aligned} C_1'(x) \cdot 1 + C_2'(x) \cos x + C_3'(x) \sin x &= 0, \\ -C_2'(x) \sin x + C_3'(x) \cos x &= 0, \\ -C_2'(x) \cos x - C_3'(x) \sin x &= \frac{1}{\sin x}. \end{aligned}$$

Решая эту систему, находим:

$$C_1'(x) = \frac{1}{\sin x}, \quad C_2'(x) = -\frac{\cos x}{\sin x}, \quad C_3'(x) = -1.$$

Интегрируем полученные соотношения:

$$\begin{aligned} C_1(x) &= \int \frac{dx}{\sin x} + C_1 = \ln \left| \operatorname{tg} \frac{x}{2} \right| + C_1, \\ C_2(x) &= \int -\frac{\cos x}{\sin x} dx + C_2 = -\ln |\sin x| + C_2, \\ C_3(x) &= -x + C_3, \end{aligned}$$

где C_1, C_2, C_3 — произвольные постоянные.

Общее решение данного уравнения в области σ запишется в виде:

$$y = \ln \left| \operatorname{tg} \frac{x}{2} \right| - \cos x \ln |\sin x| - x \sin x + C_1 + C_2 \cos x + C_3 \sin x.$$

§ 19. Линейные неоднородные уравнения высшего порядка с постоянными коэффициентами

Рассмотрим линейное неоднородное уравнение

$$L(y) = y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y = f(x) \quad (1)$$

с постоянными действительными коэффициентами и непрерывной правой частью $f(x)$ на интервале (a, b) .

Уравнение (1) всегда может быть проинтегрировано в квадратурах методом вариации постоянных, так как соответствующее однородное уравнение с постоянными коэффициентами

$$L(y) = 0 \quad (2)$$

интегрируется в элементарных функциях.

Остановимся на ряде случаев, когда уравнение (1) может быть проинтегрировано без квадратур, чисто алгебраическими методами.

Оказывается, что если функция $f(x)$ — правая часть уравнения (1) — имеет специальный вид, то частное решение уравнения (1) может быть найдено *методом неопределенных коэффициентов*.

$$f(x) = P(x)e^{ax},$$
$$y = x^r e^{ax} Q(x),$$

Доказательство.

$$F(k) = 0$$
$$F(\alpha) \neq 0.$$
$$y = Q(x) e^{ax},$$
$$L(Q(x)e^{ax}) = P(x)e^{ax}.$$
$$\begin{aligned} L(Q(x)e^{ax}) &= L[(B_0x^m + B_1x^{m-1} + \dots + B_m)e^{ax}] = \\ &= B_0L(x^me^{ax}) + B_1L(x^{m-1}e^{ax}) + \dots + B_mL(e^{ax}) = \\ &= B_0e^{ax}[x^mF'(a) + mx^{m-1}F''(a) + C_m^2x^{m-2}F'''(a) + \\ &+ \dots + F^{(m)}(a)] + B_1e^{ax}[x^{m-1}F'(a) + (m-1)x^{m-2}F''(a) + \\ &+ \dots + F^{(m-1)}(a)] + \dots + B_me^{ax}F(a) \equiv \\ &\equiv e^{ax}[A_0x^m + A_1x^{m-1} + \dots + A_m]. \end{aligned}$$
[illegible]

2. Пусть теперь число α — корень характеристического уравнения $F(k)=0$ кратности r .

$$L(Q(x)e^{ax}) \not\equiv e^{ax}P(x)$$

¹⁾ Короче, r — кратность числа α как корня характеристического уравнения.

Подсчитывая $y' = (2ax + ax^2) e^x$, $y'' = (2a + 4ax + ax^2) e^x$ и подставляя y , y' и y'' в данное уравнение, получим (после сокращения на e^x):

$$2a + 4ax + ax^2 - 2(2ax + ax^2) + ax^2 \equiv 3$$

$$\begin{array}{l|l} x^2 & a - 2a + a = 0, \\ x & 4a - 4a = 0, \\ x^0 & 2a = 3, \quad a = \frac{3}{2}. \end{array}$$

Искомое частное решение: $y = \frac{3}{2} x^2 e^x$. Общее решение данного уравнения можно записать в виде:

$$y = C_1 e^x + C_2 x e^x + \frac{3}{2} x^2 e^x,$$

где C_1 , C_2 — произвольные постоянные.

Если $a = 0$, т. е. если уравнение (1) имеет вид $L(y) = P(x)$, то частное решение уравнения находится в виде $y = x^r Q(x)$, где r — кратность нуля как корня характеристического уравнения.

Пример 3. Уравнение $y''' - y' = x^2 - 1$ имеет числа $k_1 = 0$, $k_{2,3} = \pm 1$ корнями характеристического уравнения.

Частное решение этого уравнения можно найти в виде:

$$y = x(ax^2 + bx + c) \quad (m=2, r=1).$$

Действительно, если $y = ax^3 + bx^2 + cx$, то

$$\begin{aligned} y' &= 3ax^2 + 2bx + c, \quad y'' = 6ax + 2b, \quad y''' = 6a, \\ 6a - 3ax^2 - 2bx - c &\equiv x^2 - 1, \\ \begin{array}{l|l} x^2 & -3a = 1, \quad a = -\frac{1}{3}, \\ x & -2b = 0, \quad b = 0, \\ x^0 & 6a - c = -1, \quad c = -1. \end{array} \end{aligned}$$

Искомое частное решение:

$$y = -\frac{1}{3} x^3 - x.$$

Примечания: 1. Если правая часть уравнения (1) имеет вид $P(x)e^{\alpha x}$, где $P(x)$ — многочлен с комплексными коэффициентами и α — мнимое число, то, как следует из проведенных выше рассуждений, частное решение уравнения (1) может быть найдено в виде $y = x^r e^{\alpha x} Q(x)$, где $Q(x)$ — многочлен (с комплексными коэффициентами) той же степени, что и $P(x)$, r — кратность числа α как корня характеристического уравнения.

2. Применяя теорему (см. § 18) о частном решении уравнения (1) с комбинированной правой частью, заключаем, что частное решение уравнения вида:

$$L(y) = \sum_{i=1}^N P_i(x) e^{\alpha_i x},$$

где $P_1(x), \dots, P_N(x)$ — многочлены, может быть найдено методом подбора как сумма частных решений вспомогательных уравнений

$$L(y) = P_i(x) e^{\alpha_i x}, \quad i = 1, 2, \dots, N.$$

Так, например, для отыскания частного решения уравнения $y'' + y = x^3 - 1 + x e^{-2x}$ достаточно рассмотреть уравнения $y'' + y = x^3 - 1$ и $y'' + y = x e^{-2x}$, подо-
брать их частные решения:

$$\begin{aligned} y'' + y &= x^3 - 1, \quad k_{1,2} = \pm i, \quad m=3, \quad \alpha=0, \quad r=0, \quad y_1 = ax^3 + bx^2 + cx + d, \\ y'' + y &= x e^{-2x}, \quad k_{1,2} = \pm i, \quad m=1, \quad \alpha=-2, \quad r=0, \quad y_2 = (ex + f) e^{-2x} \end{aligned}$$

(отыскание неопределенных коэффициентов проводится указанным выше методом) и сложить результаты.

Частное решение данного уравнения имеет вид:

$$y = ax^3 + bx^2 + cx + d + (ex + f) e^{-2x}.$$

Теорема 2. Если правая часть уравнения (1) имеет вид:

$$f(x) = e^{\alpha x} [P_1(x) \cos \beta x + P_2(x) \sin \beta x], \quad (3)$$

где $P_1(x)$ и $P_2(x)$ — многочлены степени соответственно m_1 и m_2 , то частное решение уравнения (1) может быть найдено в виде:

$$y = x^r e^{\alpha x} [Q_1(x) \cos \beta x + Q_2(x) \sin \beta x]. \quad (4)$$

Здесь $Q_1(x)$ и $Q_2(x)$ — многочлены степени не выше m , где m — большая из степеней m_1 , m_2 , а число r показывает, сколько раз $\alpha + i\beta$ является корнем характеристического уравнения для уравнения (2).

Приведем схему доказательства.

Так как

$$\cos \beta x = \frac{e^{i\beta x} + e^{-i\beta x}}{2}, \quad \sin \beta x = \frac{e^{i\beta x} - e^{-i\beta x}}{2i},$$

то правую часть уравнения (1):

$$L(y) = f(x)$$

можно записать в комплексной форме:

$$\tilde{P}_1(x) e^{(\alpha+i\beta)x} + \tilde{P}_2(x) e^{(\alpha-i\beta)x},$$

где $\tilde{P}_1(x)$ и $\tilde{P}_2(x)$ — многочлены (с комплексными коэффициентами) степени m , равной большей из степеней m_1 и m_2 .

По доказанному, частное решение такого уравнения (с комбинированной правой частью) можно найти в виде:

$$x^r [\tilde{Q}_1(x) e^{(\alpha+i\beta)x} + \tilde{Q}_2(x) e^{(\alpha-i\beta)x}],$$

где $\tilde{Q}_1(x)$ и $\tilde{Q}_2(x)$ — многочлены степени m , а число r равно кратности каждого из чисел $\alpha \pm i\beta$ как корня характеристического уравнения $L(y) = 0$.

Подставляя вместо функций $e^{(\alpha+i\beta)x}$ и $e^{(\alpha-i\beta)x}$ их значения: $e^{\alpha x}(\cos \beta x + i \sin \beta x)$ и $e^{\alpha x}(\cos \beta x - i \sin \beta x)$, после приведения подобных членов получим, что частное решение данного уравнения можно найти в виде:

$$y = x^r e^{\alpha x} [Q_1(x) \cos \beta x + Q_2(x) \sin \beta x],$$

где $Q_1(x)$ и $Q_2(x)$ — многочлены степени не выше m (с действительными коэффициентами). При этом степень хотя бы одного из этих многочленов будет равна m .

Примечание. Если правая часть уравнения (1) имеет вид: $e^{\alpha x} P_1(x) \cos \beta x$ (или $e^{\alpha x} P_2(x) \sin \beta x$), то, как следует из приведенных выше рассуждений, частное решение надо искать в «полном виде»:

$$x^r e^{\alpha x} [Q_1(x) \cos \beta x + Q_2(x) \sin \beta x].$$

При нахождении частного решения уравнения (1) с правой частью вида (3) можно использовать *метод неопределенных коэффициентов*. Для этого на основании теоремы 2 достаточно записать частное решение уравнения в виде (4), взяв в качестве $Q_1(x)$ и $Q_2(x)$ многочлены степени m с неопределенными коэффициентами, подставить это решение в уравнение (1) и после сокращения на $e^{\alpha x}$ приравнять коэффициенты при подобных членах. Из полученной при этом системы уравнений определяются неопределенные коэффициенты.

Пример 1. $y'' - y = 3e^{2x} \cos x$. Найти частное решение. Здесь $k_{1,2} = \pm 1$ — корни характеристического уравнения, $\alpha = 2$, $\beta = 1$, $\alpha + i\beta = 2 + i$, $r = 0$ (число $2 + i$ не является корнем характеристического уравнения) $m_1 = m_2 = 0$. Частное решение ищем в виде: $y = e^{2x}(a \cos x + b \sin x)$.

При этом

$$\begin{aligned} y' &= [(2a + b) \cos x + (2b - a) \sin x] e^{2x}, \\ y'' &= [(4b + 3a) \cos x + (-4a + 3b) \sin x] e^{2x}. \end{aligned}$$

Подставляя в уравнение y , y' , y'' и сокращая на e^{2x} , получаем:

$$\begin{aligned} (4b + 3a) \cos x + (-4a + 3b) \sin x - a \cos x - b \sin x &\equiv 3 \cos x. \\ \cos x \mid 4b + 3a - a &= 3, \quad 4b + 2a = 3, \\ \sin x \mid -4a + 3b - b &= 0, \quad 2b - 4a = 0, \\ a &= \frac{3}{10}, \quad b = \frac{6}{10}. \end{aligned}$$

Частное решение:

$$y = \frac{3}{10} e^{2x} (\cos x + 2 \sin x).$$

Пример 2. Установим вид частного решения уравнения

$$y'' - 4y' + 13y = e^{2x} (x^2 \cos 3x + \sin 3x).$$

Характеристическое уравнение $k^2 - 4k + 13 = 0$ имеет корни $k_{1,2} = 2 \pm 3i$. Правая часть уравнения удовлетворяет условиям теоремы 2. При этом $\alpha = 2$, $\beta = 3$, $r = 1$ (так как число $2 + 3i$ один раз встречается среди корней характеристического уравнения), $m_1 = 2$, $m_2 = 0$, $m = 2$.

Частное решение данного уравнения можно найти в виде:

$$y = xe^{2x} [(ax^2 + bx + c) \cos 3x + (lx^2 + mx + n) \sin 3x].$$

На отыскании неопределенных коэффициентов останавливаться не будем.

Примечание. Если правая часть уравнения (1) есть линейная комбинация функций, удовлетворяющих условиям теорем 1 или 2, то частное решение уравнения можно определить как сумму частных решений вспомогательных уравнений рассмотренного выше вида.

Например, частное решение уравнения $y'' + y = 1 + xe^{-x} + \cos x + x \sin 2x$ можно получить, сложив частные решения

$$y_1 = a, \quad y_2 = (bx + c) e^{-x}, \quad y_3 = x(m \cos x + l \sin x) \\ \text{и} \quad y_4 = (px + q) \sin 2x + (sx + t) \cos 2x$$

вспомогательных уравнений

$$y'' + y = 1, \quad y'' + y = xe^{-x}, \quad y'' + y = \cos x, \quad y'' + y = x \sin 2x.$$

§ 20. Применение теории линейных дифференциальных уравнений к исследованию некоторых колебаний

Рассмотрим ряд задач из теории колебаний, приводящихся к линейным дифференциальным уравнениям второго порядка с постоянными коэффициентами.

1. **Задача о малых колебаниях плоского математического маятника.** Материальная точка массы m подвешена на нерастяжимой невесомой нити длины l . Если нить с грузом (маятник) отклонить от положения равновесия на угол φ_0 и сообщить ей начальную скорость φ_0 , так чтобы вектор скорости $\vec{v}_0$ лежал в плоскости, проходящей через новое положение нити и положение равновесия, то маятник начнет колебаться, оставаясь в указанной плоскости. В этом случае положение маятника можно характеризовать величиной угла φ , показывающего отклонение маятника от положения равновесия. Найдем закон движения маятника, т. е. найдем закон, позволяющий определять положение маятника (угол φ) в любой момент времени t .

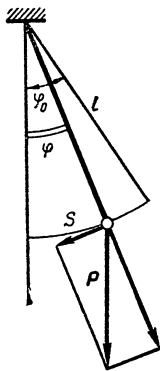


Рис. 345.

В условиях задачи материальная точка m движется по дуге окружности радиуса l с центром в точке подвеса нити. Обозначим через s величину дуги этой окружности, соответствующую углу φ : $s = l\varphi$ (рис. 345). В каждый момент времени t на точку m действуют сила тяжести $P = mg$ (ее касательная составляющая равна по величине $mg \sin \varphi$ и направлена в сторону убывания $|s|$), сила натяжения нити T , направленная вдоль нити от точки m к точке подвеса нити (ее касательная составляющая равна нулю) и сила сопротивления среды, направленная по касательной в сторону, противоположную вектору скорости движения и пропорциональная по величине этой скорости, т. е. равная $\mu \frac{ds}{dt}$, где μ — коэффициент пропорциональности, $\mu > 0$. На основании второго закона Ньютона получаем дифференциальное уравнение движения маятника:

$$m \frac{d^2 s}{dt^2} = -mg \sin \varphi - \mu \frac{ds}{dt}.$$

Подставляя значение s ($s = l\varphi$) и обозначая $\sqrt{\frac{g}{l}} = \omega$, $\frac{\mu}{m} = 2\lambda$, запишем уравнение движения маятника в виде:

$$\frac{d^2 \varphi}{dt^2} + 2\lambda \frac{d\varphi}{dt} + \omega^2 \sin \varphi = 0. \quad (1)$$

Это дифференциальное уравнение называется *уравнением свободных колебаний маятника*.

В случае, если на материальную точку m действует внешняя возмущающая сила, величина которой зависит от времени, в уравнении движения маятника появляется свободный член:

$$\frac{d^2\varphi}{dt^2} + 2\lambda \frac{d\varphi}{dt} + \omega^2 \sin \varphi = f(t). \quad (2)$$

Это уравнение называется *уравнением вынужденных колебаний маятника*. Дифференциальные уравнения (1) и (2) не являются линейными, так как неизвестная функция φ находится под знаком синуса. Однако, если рассматривать малые колебания, то $\sin \varphi$ можно заменить на φ ¹⁾.

Получаемые при этом уравнения

$$\frac{d^2\varphi}{dt^2} + 2\lambda \frac{d\varphi}{dt} + \omega^2 \varphi = 0 \quad (3)$$

и

$$\frac{d^2\varphi}{dt^2} + 2\lambda \frac{d\varphi}{dt} + \omega^2 \varphi = f(t) \quad (4)$$

называются *дифференциальными уравнениями малых колебаний маятника*.

Решения уравнений (3) и (4) дадут неточные, но практически вполне удовлетворительные приближенные значения отклонения маятника φ как функции времени t .

2. Задача о колебании груза, подвешенного на вертикальной пружине. Груз весом P подвешен на вертикальной пружине, длина которой в ненагруженном состоянии равна l . Если пружину с грузом оттянуть или сжать (в вертикальном направлении) сообщив грузу некоторую начальную скорость v_0 ²⁾, то он начнет совершать колебательные движения. Выясним, по какому закону будут происходить эти колебания, т. е. найдем формулу, выражающую отклонение груза от положения равновесия (точки O) в любой момент времени t , при условии, что масса пружины не учитывается.

Направим ось OY вертикально вверх (от точки O к точке подвеса пружины). Пусть l_0 — статическое удлинение пружины, т. е. расстояние от конца нерастянутой пружины до положения равновесия груза, y_0 — начальное отклонение, а y — отклонение груза от положения равновесия в момент времени t (рис. 346).

Действующая на груз сила равна сумме силы тяжести груза, направленной в отрицательном направлении и равной по величине P , силы сопротивления среды, которая направлена в сторону, противоположную движению груза и по величине пропорциональна скорости движения $\frac{dy}{dt}$, и силы натя-

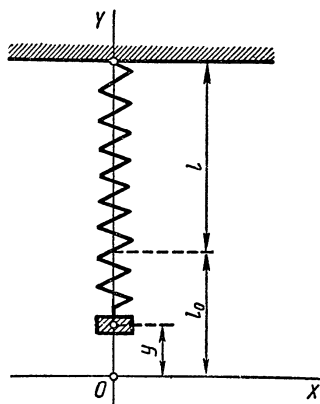


Рис. 346.

жения пружины. Сила натяжения пружины направлена по оси OY , ее величина, по закону Гука, пропорциональна удлинению пружины, т. е. равна $c(l_0 - y)$, где c — постоянный коэффициент жесткости пружины.

На основании второго закона Ньютона получаем уравнение движения груза:

$$m \frac{d^2y}{dt^2} = c(l_0 - y) - P - \mu \frac{dy}{dt},$$

где $m = \frac{P}{g}$ — масса груза, μ — коэффициент пропорциональности, $\mu > 0$. Так как в положении равновесия груза ($y = 0$) вес пружины уравновешивается силой натяжения, то

$$P = cl_0$$

и, следовательно, уравнение движения груза запишется в виде:

$$m \frac{d^2y}{dt^2} = -cy - \mu \frac{dy}{dt}$$

или, если обозначить $\sqrt{\frac{c}{m}} = \omega$, $\frac{\mu}{m} = 2\lambda$, в виде:

$$\frac{d^2y}{dt^2} + 2\lambda \frac{dy}{dt} + \omega^2 y = 0. \quad (5)$$

¹⁾ При стремлении отклонения φ к нулю $\sin \varphi$ и φ будут эквивалентными бесконечно малыми.

²⁾ Предполагается, что вектор скорости также направлен по вертикали.

Это уравнение называется *уравнением свободных колебаний* груза, подвешенного на пружине.

В случае, если на пружину с грузом действует вертикально направленная внешняя сила, величина которой зависит от времени (например, сила, заставляющая периодически колебаться точку подвеса пружины), в уравнении движения груза появится свободный член $f(t)$:

$$\frac{d^2y}{dt^2} + 2\lambda \frac{dy}{dt} + \omega^2 y = f(t). \quad (6)$$

Это линейное неоднородное дифференциальное уравнение называется *уравнением вынужденных колебаний* груза.

3. Задача о колебаниях в электрической цепи. Электродвижущая сила E включается в некоторый момент времени ($t=0$) в контур, состоящий из последовательно соединенных катушки самоиндукции L , сопротивления R и емкости C (рис. 347).

Требуется найти величину тока I в цепи как функцию времени t , если в начальный момент времени величина тока в контуре и заряд конденсатора равны нулю.

По закону Кирхгофа электродвижущая сила в цепи равна сумме падений напряжений на индуктивном (u_L) и активном (u_R) сопротивлениях и емкости (u_C):

$$E = u_L + u_R + u_C,$$

причем эти падения связаны с величиной тока I соотношениями:

$$u_L = L \frac{dI}{dt}, \quad u_R = RI, \quad u_C = \frac{1}{C} \int_0^t I(t) dt.$$

Таким образом,

$$E = L \frac{dI}{dt} + RI + \frac{1}{C} \int_0^t I(t) dt,$$

откуда дифференцированием по t получается линейное дифференциальное уравнение относительно неизвестной функции I :

$$L \frac{d^2I}{dt^2} + R \frac{dI}{dt} + \frac{1}{C} I = \frac{dE}{dt}.$$

Обозначая $\frac{R}{L} = 2\lambda$, $\sqrt{\frac{1}{CL}} = \omega$, запишем найденное уравнение в виде:

$$\frac{d^2I}{dt^2} + 2\lambda \frac{dI}{dt} + \omega^2 I = \frac{dE}{dt}.$$

В случае постоянной электродвижущей силы ($E = \text{const}$) решение задачи об электрических колебаниях сводится к решению линейного однородного дифференциального уравнения:

$$\frac{d^2I}{dt^2} + 2\lambda \frac{dI}{dt} + \omega^2 I = 0. \quad (7)$$

Если электродвижущая сила E переменная, определение переменной величины тока в цепи сведется к решению линейного неоднородного дифференциального уравнения:

$$\frac{d^2I}{dt^2} + 2\lambda \frac{dI}{dt} + \omega^2 I = f(t), \quad (8)$$

где

$$f(t) = \frac{dE}{dt}.$$

Ввиду того что полученные при решении задач 1, 2 и 3 уравнения (3) — (8) отличаются между собой только обозначениями и конкретным физическим смыслом параметров и переменных величин, для решения каждой из этих задач достаточно решить уравнение:

$$\frac{d^2y}{dx^2} + 2\lambda \frac{dy}{dx} + \omega^2 y = f(x)$$

с начальными условиями $x_0 = 0$, y_0 , y'_0 в общем виде, отвлекаясь от конкретного смысла переменных x и y и параметров λ ($\lambda \geq 0$) и ω .

Будем называть уравнение

$$\frac{d^2y}{dx^2} + 2\lambda \frac{dy}{dx} + \omega^2 y = 0 \quad (9)$$

уравнением свободных колебаний, уравнение

$$\frac{d^2y}{dx^2} + 2\lambda \frac{dy}{dx} + \omega^2 y = f(x), \quad (10)$$

где $f(x) \neq 0$, уравнением вынужденных колебаний, а решения этих уравнений — соответственно свободными или вынужденными колебаниями, или, короче, движениями. Решим эти уравнения и проведем некоторый анализ найденных решений.

Свободные колебания. Предположим сначала, что в уравнении (9) коэффициент λ равен нулю, т. е. уравнение имеет вид:

$$\frac{d^2y}{dx^2} + \omega^2 y = 0. \quad (11)$$

Этот случай соответствует отсутствию силы сопротивления среды в задачах 1 и 2 и отсутствию сопротивления R в цепи в задаче 3.

Уравнение (11) является линейным однородным уравнением второго порядка с постоянными коэффициентами, его характеристическое уравнение $k^2 + \omega^2 = 0$ имеет

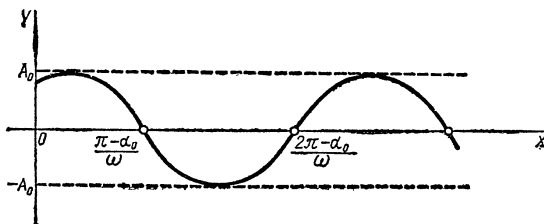


Рис. 348.

корни $k = \pm \omega i$. Общее решение уравнения (11) (в области изменения x и y) представляет собой простую гармонику:

$$y = C_1 \cos \omega x + C_2 \sin \omega x,$$

где C_1 и C_2 — произвольные постоянные.

Полагая $\sqrt{C_1^2 + C_2^2} = A$, $\sin \alpha = \frac{C_2}{A}$, $\cos \alpha = \frac{C_1}{A}$, запишем эту гармонику в виде:

$$y = A \sin(\omega x + \alpha).$$

Постоянные C_1 и C_2 (соответственно A и α) определяются из начальных условий движения:

$$C_{10} = y_0, \quad C_{20} = \frac{y'_0}{\omega}, \quad A_0 = \sqrt{y_0^2 + \frac{y'^2_0}{\omega^2}}, \quad \sin \alpha_0 = \frac{y_0}{A_0}, \quad \cos \alpha_0 = \frac{y'_0}{A_0 \omega}.$$

Таким образом, в рассматриваемом случае движение представляет собой гармоническое колебание с амплитудой A_0 , начальной фазой α_0 , частотой колебания ω и периодом колебания $T = \frac{2\pi}{\omega}$ (рис. 348).

Частота колебания и период от начальных условий движения не зависят, они определяются параметрами колеблющейся системы¹⁾. Амплитуда колебания A_0 и начальная фаза α_0 , наоборот, зависят от начальных условий движения.

Пусть теперь в уравнении свободных колебаний (9) коэффициент λ больше нуля. Этот случай соответствует учету силы сопротивления среды в задачах 1 и 2 и сопротивления цепи в задаче 3. Составляя характеристическое уравнение

$$k^2 + 2\lambda k + \omega^2 = 0,$$

находим его корни

$$k_{1,2} = -\lambda \pm \sqrt{\lambda^2 - \omega^2}.$$

Если подкоренное выражение $\lambda^2 - \omega^2$ отрицательно (сопротивление среды в задачах 1 и 2 или сопротивление цепи в задаче 3 сравнительно невелики), то корни характеристического уравнения

$$k_{1,2} = -\lambda \pm i \sqrt{\omega^2 - \lambda^2}$$

¹⁾ Будем называть их *собственной частотой* и *периодом колебания* системы.

будут комплексно сопряженными и общее решение уравнения (9) в области изменения x и y запишется в виде:

$$y = e^{-\lambda x} (C_1 \cos \sqrt{\omega^2 - \lambda^2} x + C_2 \sin \sqrt{\omega^2 - \lambda^2} x),$$

где C_1 и C_2 — произвольные постоянные.

Обозначив $A = \sqrt{C_1^2 + C_2^2}$, $\sin \alpha = \frac{C_2}{A}$, $\cos \alpha = \frac{C_1}{A}$, $\omega_1 = \sqrt{\omega^2 - \lambda^2}$, получим решение уравнения (9) в форме:

$$y = Ae^{-\lambda x} \sin (\omega_1 x + \alpha).$$

Постоянные C_1 , C_2 , а значит, A и α , определяются с помощью начальных условий.

В этом случае свободное колебание представляет собой затухающее колебание с амплитудой $Ae^{-\lambda x}$, зависящей от x , монотонно убывающей с ростом x и стремящейся к нулю при неограниченном возрастании x (рис. 349). Частота колебания $\omega_1 = \sqrt{\omega^2 - \lambda^2}$ меньше частоты колебания, соответствующей случаю $\lambda = 0$, но по-прежнему определяется самой колеблющейся системой и не зависит от начальных условий движе-

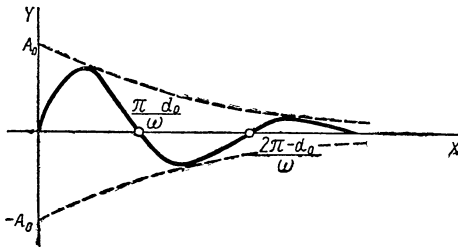


Рис. 349.



Рис. 350.

ния (собственная частота). Период колебания $T = \frac{2\pi}{\omega_1}$ также не зависит от начальных условий движения.

П р и м е ч а н и е. Рассмотренный выше случай $\lambda = 0$ получается как предельный из общего при $\lambda \rightarrow 0$.

Если подкоренное выражение $\lambda^2 - \omega^2$ положительно (сопротивление среды в задачах 1 и 2 и сопротивление цепи в задаче 3 велико), то корни

$$k_{1,2} = -\lambda \pm \sqrt{\lambda^2 - \omega^2}$$

характеристического уравнения действительны, различны и отрицательны, решение уравнения (9) запишется в виде:

$$y = C_1 e^{-(\lambda - \sqrt{\lambda^2 - \omega^2})x} + C_2 e^{-(\lambda + \sqrt{\lambda^2 - \omega^2})x},$$

где C_1 и C_2 — постоянные, определяемые с помощью начальных условий.

В этом случае движение носит неперiodический характер, причем величина y стремится к нулю при неограниченном возрастании x (рис. 350).

Если $\lambda^2 - \omega^2 = 0$, то $k_{1,2} = -\lambda$, т. е. решение уравнения (9) дается равенством:

$$y = C_1 e^{-\lambda x} + C_2 x e^{-\lambda x}.$$

Движение снова неперiodическое, величина y при неограниченном возрастании x стремится к нулю.

2. Вынужденные колебания. Рассмотрим уравнение вынужденных колебаний (10) в случае периодической внешней возмущающей силы, предположив для простоты, что $\lambda = 0$:

$$\frac{d^2 y}{dx^2} + \omega^2 y = B \sin \beta x. \quad (12)$$

Это линейное неоднородное дифференциальное уравнение с постоянными коэффициентами и правой частью специального вида.

Характеристическое уравнение имеет корни $k_{1,2} = \pm \omega i$, следовательно, общее решение соответствующего однородного уравнения, т. е. уравнения свободных колеба-

ний: $\frac{d^2 y}{dx^2} + \omega^2 y = 0$, дается равенством:

$$y = A \sin (\omega x + \alpha),$$

где A и α — произвольные постоянные.

Для нахождения частного решения неоднородного уравнения (12) рассмотрим отдельно два возможных случая.

1) Пусть $\omega \neq \beta$, т. е. частота свободных (собственных) колебаний системы не совпадает с частотой возмущающей силы. Так как число $i\beta$ не является корнем характеристического уравнения, частное решение уравнения (12) можно найти в виде простой гармоник:

$$y_{\text{ч}} = a \cos \beta x + b \sin \beta x.$$

При этом $y'_{\text{ч}} = -a\beta \sin \beta x + b\beta \cos \beta x$, $y''_{\text{ч}} = -a\beta^2 \cos \beta x - b\beta^2 \sin \beta x$. Подставляя $y_{\text{ч}}$, $y'_{\text{ч}}$ и $y''_{\text{ч}}$ в уравнение (12), получаем:

$$a = 0, \quad b = \frac{B}{\omega^2 - \beta^2}.$$

Таким образом, $y_{\text{ч}} = \frac{B}{\omega^2 - \beta^2} \sin \beta x$.

Это частное решение определяет вынужденное колебание, создаваемое возмущающей силой. Вынужденное колебание имеет ту же частоту, что и возмущающая сила. Амплитуда этого колебания постоянна, но в случае близости частот ω и β может быть велика даже при малой амплитуде возмущающей силы ¹⁾. Общее решение уравнения (12) в данном случае имеет вид:

$$y = A \sin(\omega x + \alpha) + \frac{B}{\omega^2 - \beta^2} \sin \beta x.$$

Для конкретного движения числа A и α определяются с помощью начальных условий движения.

В данном случае мы имеем сложное колебательное движение, получающееся от сложения собственных свободных колебаний системы и колебаний, обусловленных действием возмущающей силы.

2) Пусть теперь $\omega = \beta$, т. е. частота собственных свободных колебаний системы совпадает с частотой возмущающей внешней силы.

Так как число $i\beta = i\omega$ является простым корнем характеристического уравнения $k^2 + \omega^2 = 0$, то частное решение уравнения

$$\frac{d^2 y}{dx^2} + \omega^2 y = B \sin \omega x \quad (13)$$

будем искать в виде:

$$y_{\text{ч}} = x (a \cos \omega x + b \sin \omega x).$$

Подсчитывая $y'_{\text{ч}}$ и $y''_{\text{ч}}$ и подставляя $y_{\text{ч}}$, $y'_{\text{ч}}$ и $y''_{\text{ч}}$ в уравнение (13), находим:

$$a = -\frac{B}{2\omega}, \quad b = 0,$$

т. е.

$$y_{\text{ч}} = -\frac{B}{2\omega} x \cos \omega x.$$

Общее решение уравнения (12) в этом случае имеет форму:

$$y = A \sin(\omega x + \alpha) - \frac{B}{2\omega} x \cos \omega x. \quad (14)$$

Постоянные A и α определяются из начальных условий.

Из соотношения (14) следует, что рассматриваемое движение снова складывается из двух колебательных движений — периодического колебания с частотой ω (свободного колебания) и колебания с той же частотой ω и возрастающей с ростом x по формуле $\frac{B}{2\omega} x$ амплитудой (вынужденного колебания). Даже в случае, если амплитуда внешней силы мала, амплитуда вынужденного колебания станет сколь угодно большой для достаточно больших значений x . Это явление неограниченного возрастания амплитуды колебания под влиянием даже малых периодических возмущающих сил называется в технике явлением *резонанса*. Наличие резонанса может привести к разрушению колеблющейся системы.

П р и м е ч а н и е. Явление резонанса в рассмотренных выше примерах можно проследить на задачах 2 и 3. В задаче 1 уравнение (4) описывает действительные колебания маятника только для малых отклонений маятника.

Исследование решений уравнения (10) в случае $\lambda > 0$ проводится аналогично изученным случаям.

Предоставляем это сделать читателю.

¹⁾ Возможностью создания больших колебаний часто пользуются в различного рода усилителях.

$$\begin{aligned} F_1(x, y_1, y_2, \dots, y_n, y'_1, y'_2, \dots, y'_n) &= 0, \\ F_2(x, y_1, y_2, \dots, y_n, y'_1, y'_2, \dots, y'_n) &= 0, \\ &\vdots \\ F_n(x, y_1, y_2, \dots, y_n, y'_1, y'_2, \dots, y'_n) &= 0, \end{aligned} \quad (1)$$
$$\begin{aligned} z' - \frac{1}{3}(4-x)y - z + 2x &= 0, \\ x^2 y' - 2z' - 6z + 4x &= 0 \end{aligned} \quad (2)$$
$$\begin{aligned} 3x^2 - 2x - \frac{1}{3}(4-x)3x^2 - x^3 + x^2 + 2x &\equiv 0, \\ 6x^3 - 2(3x^2 - 2x) - 6(x^3 - x^2) + 4x &\equiv 0. \end{aligned}$$
$$\begin{aligned} \frac{dy_1}{dx} &= f_1(x, y_1, y_2, \dots, y_n), \\ \frac{dy_2}{dx} &= f_2(x, y_1, y_2, \dots, y_n), \\ &\vdots \\ \frac{dy_n}{dx} &= f_n(x, y_1, y_2, \dots, y_n). \end{aligned} \quad (3)$$
$$\frac{dy}{dx} = f(x, y). \quad (4)$$
$$\begin{aligned}\frac{dy}{dx} &= f_1(x, y, z), \\ \frac{dz}{dx} &= f_2(x, y, z),\end{aligned}\tag{5}$$

т. е. пару функций $y = \varphi_1(x)$, $z = \varphi_2(x)$, можно рассматривать как некоторую кривую в пространстве XYZ трех измерений. Эту кривую называют *интегральной кривой* системы (5) в этом пространстве.

В случае $n > 2$ решение нормальной системы (3):

$$y_1 = \varphi_1(x), y_2 = \varphi_2(x), \dots, y_n = \varphi_n(x)$$

нельзя аналогичным образом изобразить кривой, не выходя за пределы трехмерного пространства.

Однако, обобщая геометрическую терминологию, по-прежнему будем считать, что решение

$$y_1 = \varphi_1(x), y_2 = \varphi_2(x), \dots, y_n = \varphi_n(x) \quad (6)$$

системы (3) представляет собой *интегральную кривую* системы (3), лежащую в $(n+1)$ -мерном пространстве переменных $x, y_1, y_2, \dots, y_n$.

Условимся говорить, что решение (6) системы (3) удовлетворяет начальным данным $x_0, y_{10}, y_{20}, \dots, y_{n0}$, если $\varphi_1(x_0) = y_{10}, \varphi_2(x_0) = y_{20}, \dots, \varphi_n(x_0) = y_{n0}$. Геометрически это значит, что интегральная кривая (6) проходит через точку $(x_0, y_{10}, y_{20}, \dots, y_{n0})$ $(n+1)$ -мерного пространства.

Задача Коши для нормальной системы (3) и начальных данных $x_0, y_{10}, \dots, y_{n0}$ ставится следующим образом: найти решение системы (3), удовлетворяющее начальным данным $x_0, y_{10}, \dots, y_{n0}$.

Достаточные условия для существования и единственности решения задачи Коши даются следующей теоремой.

Теорема Коши. Если в некоторой области D $(n+1)$ -мерного пространства функции $f_1(x, y_1, y_2, \dots, y_n), f_2(x, y_1, y_2, \dots, y_n), \dots, f_n(x, y_1, y_2, \dots, y_n)$ непрерывны и имеют непрерывные частные производные по $y_1, y_2, \dots, y_n$, то для любой точки $(x_0, y_{10}, y_{20}, \dots, y_{n0})$ области D существует, и притом единственное, решение

$$y_1 = \varphi_1(x), y_2 = \varphi_2(x), \dots, y_n = \varphi_n(x)$$

системы

$$\frac{dy_1}{dx} = f_1(x, y_1, y_2, \dots, y_n),$$

$$\frac{dy_2}{dx} = f_2(x, y_1, y_2, \dots, y_n),$$

$$\dots \dots \dots$$

$$\frac{dy_n}{dx} = f_n(x, y_1, y_2, \dots, y_n),$$

определенное в некоторой окрестности точки x_0 и удовлетворяющее начальным условиям

$$x_0, y_{10}, y_{20}, \dots, y_{n0}.$$

Из теоремы Коши следует, что в указанной области D система (3) имеет бесчисленное множество решений. Действительно, изменяя значения $y_{10}, y_{20}, \dots, y_{n0}$ в некоторых пределах, для каждой системы чисел $x_0, y_{10}, y_{20}, \dots, y_{n0}$ получим «свое» решение системы:

$$y_1 = \varphi_1(x, y_{10}, y_{20}, \dots, y_{n0}), y_2 = \varphi_2(x, y_{10}, y_{20}, \dots, y_{n0}), \dots \\ \dots, y_n = \varphi_n(x, y_{10}, y_{20}, \dots, y_{n0}).$$

Будем называть n функций, зависящих от n произвольных постоянных $C_1, C_2, \dots, C_n$:

$$y_1 = \varphi_1(x, C_1, C_2, \dots, C_n), y_2 = \varphi_2(x, C_1, C_2, \dots, C_n), \dots \\ \dots, y_n = \varphi_n(x, C_1, C_2, \dots, C_n) \quad (7)$$

¹⁾ Задание нормальной системы уравнений (3) можно по аналогии с двухмерным случаем геометрически толковать как задание «поля направлений» в некоторой области $(n+1)$ -мерного пространства. Подробнее об этом см. в книге: И. Г. Петровский, Лекции по теории обыкновенных дифференциальных уравнений.

Решения, получающиеся из общего при конкретных значениях постоянных $C_1, C_2, \dots, C_n$, будем называть *частными решениями* системы (3).

$$\begin{array}{l} \varphi_1(x_0, C_1, C_2, \dots, C_n) = y_{10}, \\ \varphi_2(x_0, C_1, C_2, \dots, C_n) = y_{20}, \\ .\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad., \\ \varphi_n(x_0, C_1, C_2, \dots, C_n) = y_{n0}. \end{array}$$

Нормальная система уравнений (3) в случае $n=3$ допускает простую механическую интерпретацию.

При этом решение

системы

$$\begin{aligned}\frac{dx}{dt} &= f_1(t, x, y, z), \\ \frac{dy}{dt} &= f_2(t, x, y, z), \\ \frac{dz}{dt} &= f_3(t, x, y, z)\end{aligned}\tag{8}$$

Кривая пространства XYZ , задаваемая уравнениями

$$x = \varphi_1(t), \quad y = \varphi_2(t), \quad z = \varphi_3(t),$$

Таким образом, задать систему (8) — значит задать в пространстве XYZ (или некоторой его области) в каждый момент времени t компоненты векторов скорости движущихся частиц, или, короче, задать в каждый момент времени t *поле скоростей*.

При такой интерпретации система (8) называется *динамической*, пространство XYZ — *фазовым*, а решения системы (8) — *движениями*.

488

в любой момент времени в этой точке вектор скорости имеет одну и ту же величину и направление.

В случае нестационарного движения поле скоростей изменяется со временем и траектории движений могут пересекаться.

Аналогичным образом можно нормальную систему дифференциальных уравнений с любым числом n неизвестных рассматривать как *динамическую систему*, а ее решения — как *движения* в соответствующем фазовом пространстве n измерений.

Изучение нормальных систем и их свойств тесно переплетается с изучением дифференциальных уравнений высших порядков от одной неизвестной функции.

Всякое дифференциальное уравнение n -го порядка, разрешенное относительно старшей производной:

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}), \quad (9)$$

может быть заменено эквивалентной ему нормальной системой из n дифференциальных уравнений.

Действительно, обозначим

$$y = z_1, \quad y' = z_2, \quad \dots, \quad y^{(n-1)} = z_n$$

и будем рассматривать $z_1, z_2, \dots, z_n$ как новые неизвестные функции от x . Эти функции связаны между собой соотношениями:

$$\begin{aligned} \frac{dz_1}{dx} &= z_2, \\ \frac{dz_2}{dx} &= z_3, \\ &\dots \dots \dots \\ \frac{dz_{n-1}}{dx} &= z_n, \\ \frac{dz_n}{dx} &= f(x, z_1, z_2, \dots, z_n), \end{aligned} \quad (10)$$

составляющими нормальную систему дифференциальных уравнений.

Если $z_1, z_2, \dots, z_n$ — решение системы (10), то $z_1 = y$ — решение уравнения (9) и, наоборот, если y — решение уравнения (9), то функции

$$z_1 = y, \quad z_2 = y', \quad \dots, \quad z_n = y^{(n-1)}$$

удовлетворяют системе (10), т. е. составляют ее решение.

При определенных условиях справедливо и обратное утверждение: всякая нормальная система дифференциальных уравнений (3):

$$\begin{aligned} \frac{dy_1}{dx} &= f_1(x, y_1, y_2, \dots, y_n), \\ \frac{dy_2}{dx} &= f_2(x, y_1, y_2, \dots, y_n), \\ &\dots \dots \dots \\ \frac{dy_n}{dx} &= f_n(x, y_1, y_2, \dots, y_n) \end{aligned}$$

эквивалентна некоторому дифференциальному уравнению n -го порядка с одной неизвестной функцией, разрешенному относительно производной.

Действительно, дифференцируем по x обе части первого из уравнений системы:

$$\frac{d^2 y_1}{dx^2} = \frac{\partial f_1}{\partial x} + \frac{\partial f_1}{\partial y_1} \frac{dy_1}{dx} + \dots + \frac{\partial f_1}{\partial y_n} \frac{dy_n}{dx}.$$

Заменяя $\frac{dy_1}{dx}, \frac{dy_2}{dx}, \dots, \frac{dy_n}{dx}$ их выражениями через $x, y_1, \dots, y_n$, получим:

$$\frac{d^2 y_1}{dx^2} = \frac{\partial f_1}{\partial x} + \frac{\partial f_1}{\partial y_1} f_1 + \dots + \frac{\partial f_1}{\partial y_n} f_n$$

или, короче,

$$\frac{d^2 y_1}{dx^2} = \tilde{f}_2(x, y_1, y_2, \dots, y_n).$$

Дифференцируя это уравнение по x и используя снова уравнения системы, получим:

$$\frac{d^2 y_1}{dx^2} = \tilde{f}_2(x, y_1, y_2, \dots, y_n) \text{ и т. д.}$$

Таким образом, будет составлена система уравнений:

$$\begin{aligned} \frac{dy_1}{dx} &= f_1(x, y_1, y_2, \dots, y_n), \\ \frac{d^2 y_1}{dx^2} &= \tilde{f}_2(x, y_1, y_2, \dots, y_n), \\ &\dots \dots \dots \\ \frac{d^n y_1}{dx^n} &= \tilde{f}_n(x, y_1, y_2, \dots, y_n). \end{aligned}$$

Из первых $(n-1)$ соотношений неизвестные $y_2, y_3, \dots, y_n$, вообще говоря, выражаются через $x, y_1, \frac{dy_1}{dx}, \dots, \frac{d^{n-1} y_1}{dx^{n-1}}$. Подставляя их значения в n -е уравнение, мы приходим к дифференциальному уравнению n -го порядка относительно неизвестной функции y_1 :

$$\frac{d^n y_1}{dx^n} = f\left(x, y_1, \frac{dy_1}{dx}, \dots, \frac{d^{n-1} y_1}{dx^{n-1}}\right). \quad (11)$$

Если $y_1, y_2, \dots, y_n$ — решение системы (3), то y_1 удовлетворяет уравнению (11). Наоборот, если y_1 — решение уравнения (11), то, находя $y'_1, y''_1, \dots, y^{(n-1)}_1$ и подставляя эти значения в соотношения для $y_2, y_3, \dots, y_n$, найдем эти функции. Можно показать, что полученные таким образом функции $y_1, y_2, \dots, y_n$ удовлетворяют системе (3).

Таким образом, в случае, если указанная выше система уравнений разрешима относительно $y_2, y_3, \dots, y_n$, то интегрирование системы (3) сводится к интегрированию дифференциального уравнения высшего порядка с неизвестной функцией y_1 .

Аналогичные рассуждения можно провести по отношению к любой другой из функций $y_2, y_3, \dots, y_n$.

Сведение интегрирования нормальной системы дифференциальных уравнений к интегрированию одного уравнения высшего порядка является одним из основных методов интегрирования нормальных систем.

Пример. Решим систему уравнений:

$$\frac{dy}{dx} = y + z, \quad \frac{dz}{dx} = y + z + x.$$

Продифференцировав первое из уравнений по x :

$$\frac{d^2 y}{dx^2} = \frac{dy}{dx} + \frac{dz}{dx}$$

и заменив $\frac{dz}{dx}$ его значением, получим:

$$\frac{d^2 y}{dx^2} = \frac{dy}{dx} + y + z + x.$$

Подставляя вместо $y + z$ из первого уравнения $\frac{dy}{dx}$, приходим к линейному неоднородному уравнению второго порядка:

$$\frac{d^2 y}{dx^2} - 2 \frac{dy}{dx} = x. \quad (12)$$

Решаем это уравнение:

$$\frac{d^2 y}{dx^2} - 2 \frac{dy}{dx} = 0, \quad k^2 - 2k = 0, \quad k_1 = 0, \quad k_2 = 2,$$

$y = C_1 + C_2 e^{2x}$ — общее решение однородного уравнения.

$y_4 = x(ax + b) = ax^2 + bx$ — частное решение неоднородного уравнения.

$$y'_4 = 2ax + b, \quad y''_4 = 2a, \quad 2a - 2(2ax + b) = x, \quad a = -\frac{1}{4}, \quad b = -\frac{1}{4},$$

$$y_4 = -\frac{1}{4}x(x+1),$$

$$y = C_1 + C_2 e^{2x} - \frac{1}{4}x(x+1),$$

где C_1 и C_2 — произвольные постоянные — общее решение вспомогательного уравнения (12).

Подставляя найденное значение y и $y' = 2C_2 e^{2x} - \frac{x}{2} - \frac{1}{4}$ в первое уравнение системы, находим:

$$z = 2C_2 e^{2x} - \frac{x}{2} - \frac{1}{4} - C_1 - C_2 e^{2x} + \frac{1}{4} x^2 + \frac{1}{4} x,$$

т. е.

$$z = -C_1 + C_2 e^{2x} + \frac{1}{4} (x^2 - x - 1).$$

Семейство решений данной системы уравнений дается функциями:

$$y = C_1 + C_2 e^{2x} - \frac{1}{4} x (x + 1), \quad z = -C_1 + C_2 e^{2x} + \frac{1}{4} (x^2 - x - 1),$$

где C_1 и C_2 — произвольные постоянные.

Примечание. После нахождения одной из функций выражения для остальных следует искать путем дифференцирования и исключения.

§ 22. Системы линейных дифференциальных уравнений

Нормальная система дифференциальных уравнений (3) называется *линейной*, если функции $f_1, f_2, \dots, f_n$ линейны относительно неизвестных функций. Таким образом, линейная система дифференциальных уравнений имеет вид:

$$\begin{aligned} y_1' &= p_{11}(x)y_1 + p_{12}(x)y_2 + \dots + p_{1n}(x)y_n + q_1(x), \\ y_2' &= p_{21}(x)y_1 + p_{22}(x)y_2 + \dots + p_{2n}(x)y_n + q_2(x), \\ &\vdots \\ y_n' &= p_{n1}(x)y_1 + p_{n2}(x)y_2 + \dots + p_{nn}(x)y_n + q_n(x). \end{aligned} \quad (1)$$

Легко видеть, что если функции $p_{ij}(x)$, $i, j = 1, 2, \dots, n$, и $q_i(x)$, $i = 1, 2, \dots, n$, непрерывны на некотором интервале (a, b) , то правые части этих уравнений в области $\{a < x < b, -\infty < y_1 < +\infty, -\infty < y_2 < +\infty, \dots, -\infty < y_n < +\infty\}$ непрерывны и имеют непрерывные частные производные по $y_1, y_2, \dots, y_n$. Отсюда, на основании теоремы Коши, следует, что для любого значения $x = x_0$, $a < x_0 < b$, и любых n чисел $y_{10}, y_{20}, \dots, y_{n0}$ существует, и при том единственное, решение системы (1), определенное в некоторой окрестности точки x_0 и удовлетворяющее начальным условиям $x_0, y_{10}, y_{20}, \dots, y_{n0}$.

Можно показать, что это решение системы определено не только в окрестности точки x_0 , но и на всем интервале (a, b) .

Если на интервале (a, b) функции $q_1(x), q_2(x), \dots, q_n(x)$ равны тождественно нулю, то система (1) называется *линейной однородной*; в противном случае, т. е. если хоть одна из этих функций не равна тождественно нулю, система называется *линейной неоднородной*.

Так же как и для линейных уравнений высшего порядка, существует хорошо разработанная теория линейных однородных и неоднородных систем дифференциальных уравнений, изучающая свойства решений, структуру общих решений этих систем уравнений, метод вариации постоянных для неоднородных линейных систем и т. д.

Рассмотрим подробнее *линейные однородные системы дифференциальных уравнений с постоянными коэффициентами*. Для простоты записи ограничимся случаем $n = 3$.

Обозначая неизвестные функции независимой переменной x через y, z и w , запишем линейную однородную систему с постоянными действительными коэффициентами в виде:

$$\begin{aligned} \frac{dy}{dx} &= a_{11}y + a_{12}z + a_{13}w, \\ \frac{dz}{dx} &= a_{21}y + a_{22}z + a_{23}w, \\ \frac{dw}{dx} &= a_{31}y + a_{32}z + a_{33}w. \end{aligned} \quad (2)$$

Решения этой системы обладают следующими свойствами:

1. Если y, z, w — решение системы (2), то Cy, Cz, Cw , где C — любое число, также есть решение системы (2).

2. Если y_1, z_1, w_1 и y_2, z_2, w_2 — решения системы (2), то функции $y_1 + y_2, z_1 + z_2, w_1 + w_2$ также будут составлять решение системы (2).

Справедливость этих свойств устанавливается непосредственной подстановкой в уравнения системы указанных функций.

Из свойств 1) и 2) следует, что если $y_1, z_1, w_1; y_2, z_2, w_2$ и y_3, z_3, w_3 — решения системы (2), то для любых значений чисел C_1, C_2, C_3 линейная комбинация этих решений $C_1y_1 + C_2y_2 + C_3y_3, C_1z_1 + C_2z_2 + C_3z_3, C_1w_1 + C_2w_2 + C_3w_3$ также является решением системы (2)¹⁾.

Будем искать ненулевые решения системы (2) вида:

$$y = \alpha e^{kx}, \quad z = \beta e^{kx}, \quad w = \gamma e^{kx}, \quad (3)$$

где α, β, γ и k — некоторые числа, которые надо подобрать так, чтобы функции (3) удовлетворяли системе (2).

Подставляя функции (3) и их производные в уравнения системы, получим (после сокращения на e^{kx} и переноса всех членов в одну часть равенства):

$$\begin{aligned} (a_{11} - k)\alpha + a_{12}\beta + a_{13}\gamma &= 0, \\ a_{21}\alpha + (a_{22} - k)\beta + a_{23}\gamma &= 0, \\ a_{31}\alpha + a_{32}\beta + (a_{33} - k)\gamma &= 0. \end{aligned} \quad (4)$$

Систему (4) будем рассматривать как однородную алгебраическую систему из трех уравнений с тремя неизвестными α, β и γ . Для того чтобы эта система имела ненулевое решение, необходимо и достаточно, чтобы определитель системы был равен нулю. Таким образом, число k должно удовлетворять уравнению:

$$\begin{vmatrix} a_{11} - k & a_{12} & a_{13} \\ a_{21} & a_{22} - k & a_{23} \\ a_{31} & a_{32} & a_{33} - k \end{vmatrix} = 0. \quad (5)$$

Уравнение (5) называется *характеристическим* (или *вековым*) уравнением системы (2). Это — уравнение третьей степени относительно k . По основной теореме алгебры оно имеет три корня: k_1, k_2 и k_3 . Каждому из этих корней соответствует ненулевое решение системы (4) (обозначим эти решения соответственно $\alpha_1, \beta_1, \gamma_1; \alpha_2, \beta_2, \gamma_2$ и $\alpha_3, \beta_3, \gamma_3$), а значит, ненулевое решение данной системы дифференциальных уравнений (2):

$$\begin{aligned} y_1 &= \alpha_1 e^{k_1 x}, & z_1 &= \beta_1 e^{k_1 x}, & w_1 &= \gamma_1 e^{k_1 x}, \\ y_2 &= \alpha_2 e^{k_2 x}, & z_2 &= \beta_2 e^{k_2 x}, & w_2 &= \gamma_2 e^{k_2 x}, \\ y_3 &= \alpha_3 e^{k_3 x}, & z_3 &= \beta_3 e^{k_3 x}, & w_3 &= \gamma_3 e^{k_3 x}. \end{aligned}$$

Линейная комбинация этих решений с произвольными коэффициентами:

$$\begin{aligned} y &= C_1 \alpha_1 e^{k_1 x} + C_2 \alpha_2 e^{k_2 x} + C_3 \alpha_3 e^{k_3 x}, & z &= C_1 \beta_1 e^{k_1 x} + C_2 \beta_2 e^{k_2 x} + C_3 \beta_3 e^{k_3 x}, \\ w &= C_1 \gamma_1 e^{k_1 x} + C_2 \gamma_2 e^{k_2 x} + C_3 \gamma_3 e^{k_3 x} \end{aligned}$$

(по доказанному выше) также будет решением системы (2).

Если корни k_1, k_2 и k_3 характеристического уравнения различны, то можно показать, что это решение будет общим решением системы (2) в области

$$\{-\infty < x < +\infty, -\infty < y_1 < +\infty, \dots, -\infty < y_n < +\infty\}.$$

¹⁾ Свойства 1), 2) и указанное следствие, как легко видеть, справедливы и для линейной однородной системы с переменными коэффициентами.

Примечание. Так же как и для одного линейного уравнения, устанавливается, что если

$$y = u + iv, \quad z = s + it, \quad w = p + iq$$

— комплексное решение системы (2) с действительными коэффициентами, то действительные и мнимые части этих функций, т. е. u, s, p и v, t, q , также составляют решения системы (2).

В случае, если $k_{1,2} = \alpha \pm i\beta$ — корни характеристического уравнения, то найденные указанным выше методом комплексные решения можно заменить действительными решениями, отделяя действительные и мнимые части составляющих их функций так же, как это делалось на стр. 469.

В случае кратных корней характеристического уравнения отыскание общего решения системы (2) значительно сложнее. Этот вопрос в данном курсе не рассматривается.

Пример 1. Найти общее решение системы:

$$\begin{aligned} y' &= z + w, \\ z' &= 3y + w, \\ w' &= 3y + z. \end{aligned} \quad (6)$$

Ищем решение в виде: $y = ae^{kx}$, $z = \beta e^{kx}$, $w = \gamma e^{kx}$. Подставляя эти функции в уравнения и сокращая на e^{kx} , получаем систему:

$$\begin{aligned} -\alpha k + \beta + \gamma &= 0, \\ 3\alpha - \beta k + \gamma &= 0, \\ 3\alpha + \beta - \gamma k &= 0. \end{aligned} \quad (7)$$

Характеристическое уравнение системы (6):

$$\begin{vmatrix} -k & 1 & 1 \\ 3 & -k & 1 \\ 3 & 1 & -k \end{vmatrix} = 0, \text{ т. е. } k^3 - 7k - 6 = 0.$$

Корни характеристического уравнения $k_1 = -1$, $k_2 = -2$, $k_3 = 3$.

1) Подставляем $k_1 = -1$ в систему (7):

$$\begin{aligned} \alpha + \beta + \gamma &= 0, \\ 3\alpha + \beta + \gamma &= 0, \\ 3\alpha + \beta + \gamma &= 0. \end{aligned}$$

Третье уравнение отбрасываем (оно следует из первых двух), переносим γ вправо и находим α и β :

$$\alpha = 0, \quad \beta = -\gamma, \quad \gamma = \gamma, \text{ где } \gamma - \text{любое число.}$$

Полагая $\gamma = 1$, находим решение системы (6):

$$y_1 = 0, \quad z_1 = -e^{-x}, \quad w_1 = e^{-x}.$$

2) Подставляем $k_2 = -2$ в систему (7):

$$\begin{aligned} 2\alpha + \beta + \gamma &= 0, \\ 3\alpha + 2\beta + \gamma &= 0, \\ 3\alpha + \beta + 2\gamma &= 0. \end{aligned}$$

Решая систему из первых двух уравнений:

$$\begin{aligned} 2\alpha + \beta &= -\gamma, \\ 3\alpha + 2\beta &= -\gamma, \end{aligned}$$

получаем $\alpha = -\gamma$, $\beta = \gamma$, $\gamma = \gamma$, где γ — любое число. Полагая $\gamma = 1$, находим решение системы (6):

$$y_2 = -e^{-2x}, \quad z_2 = e^{-2x}, \quad w = e^{-2x}.$$

3) Подставляем $k_3 = 3$ в систему (7):

$$\begin{aligned} -3\alpha + \beta + \gamma &= 0, \\ 3\alpha - 3\beta + \gamma &= 0, \\ 3\alpha + \beta - 3\gamma &= 0. \end{aligned}$$

Отбрасываем снова третье уравнение и решаем систему

$$\begin{aligned} -3\alpha + \beta &= -\gamma, \\ 3\alpha - 3\beta &= -\gamma. \end{aligned}$$

Получаем $\beta = \gamma$, $\alpha = \frac{2}{3}\gamma$, $\gamma = \gamma$. Полагаем $\gamma = 3$. Тогда $\alpha = 2$, $\beta = 3$, $\gamma = 3$.

$$y_3 = 2e^{3x}, \quad z_3 = 3e^{3x}, \quad w_3 = 3e^{3x}.$$

Общее решение системы имеет вид:

$$y = -C_2 e^{-2x} + 2C_3 e^{3x}, \quad z = -C_1 e^{-x} + C_2 e^{-2x} + 3C_3 e^{3x}, \quad w = C_1 e^{-x} + C_2 e^{-2x} + 3C_3 e^{3x}.$$

Пример 2.

$$\frac{dy}{dx} = -7y + z,$$

$$\frac{dz}{dx} = -2y - 5z.$$

Ищем решение в виде: $y = \alpha e^{kx}$, $z = \beta e^{kx}$. При этом получаем:

$$\begin{aligned} \alpha(k+7) - \beta &= 0, \\ 2\alpha + \beta(k+5) &= 0. \end{aligned} \tag{8}$$

Характеристическое уравнение:

$$\begin{vmatrix} k+7 & -1 \\ 2 & k+5 \end{vmatrix} = 0, \text{ т. е. } k^2 + 12k + 37 = 0.$$

Корни $k_{1,2} = -6 \pm i$ — комплексно сопряженные.

1) $k_1 = -6 + i$. Подставляя в систему (8), получаем:

$$\begin{aligned} (1+i)\alpha - \beta &= 0, \\ 2\alpha + \beta(-1+i) &= 0. \end{aligned}$$

Второе уравнение — следствие первого. Из первого уравнения находим: $\beta = (1+i)\alpha$. Взяв $\alpha = 1$, имеем $\beta = 1+i$, т. е. $y_1 = e^{(-6+i)x}$, $z_1 = (1+i)e^{(-6+i)x}$ — решение данной системы.

2) $k_2 = -6 - i$. Система (8) имеет вид:

$$\begin{aligned} (1-i)\alpha - \beta &= 0, \\ 2\alpha + \beta(-1-i) &= 0, \end{aligned}$$

откуда следует, что $\beta = (1-i)\alpha$. Взяв $\alpha = 1$, имеем $\beta = 1-i$, т. е.

$$y_2 = e^{(-6-i)x}, \quad z_2 = (1-i)e^{(-6-i)x}.$$

Найдем действительные и мнимые части найденных функций:

$$\begin{aligned} y_1 &= e^{-6x} \cos x + ie^{-6x} \sin x, \quad z_1 = e^{-6x} (\cos x - \sin x) + ie^{-6x} (\sin x + \cos x), \\ y_2 &= e^{-6x} \cos x - ie^{-6x} \sin x, \quad z_2 = e^{-6x} (\cos x - \sin x) - ie^{-6x} (\sin x + \cos x). \end{aligned}$$

Действительные решения системы имеют вид:

$$\begin{aligned} \tilde{y}_1 &= e^{-6x} \cos x, \quad \tilde{z}_1 = e^{-6x} (\cos x - \sin x), \\ \tilde{y}_2 &= e^{-6x} \sin x, \quad \tilde{z}_2 = e^{-6x} (\sin x + \cos x). \end{aligned}$$

Общее решение (во всем пространстве XYZ):

$$y = e^{-6x} (C_1 \cos x + C_2 \sin x), \quad z = e^{-6x} [C_1 (\cos x - \sin x) + C_2 (\sin x + \cos x)].$$

ГЛАВА IX

ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ МНОГИХ ПЕРЕМЕННЫХ

Раздел I

ИНТЕГРАЛЫ, ЗАВИСЯЩИЕ ОТ ПАРАМЕТРОВ

§ 1. Интегралы, зависящие от параметров

Рассмотрим функцию нескольких переменных

$$u = f(x, y, z, \dots, w),$$

определенную и непрерывную в области

$$a \leq x \leq b, \quad c \leq y \leq d, \quad l \leq z \leq m, \quad \dots, \quad s \leq w \leq t.$$

Закрепив любые значения $y, z, \dots, w$, мы получим функцию одной переменной x :

$$u = \varphi(x) = f(x, y, z, \dots, w),$$

непрерывную, а значит, интегрируемую, на отрезке $a \leq x \leq b$. Интеграл от этой функции

$$\int_a^b \varphi(x) dx = \int_a^b f(x, y, z, \dots, w) dx, \quad \text{где } a \leq \alpha \leq b, \quad \alpha \leq \beta \leq b,$$

зависит от выбранных нами значений $y, z, \dots, w$, т. е. является функцией от $y, z, \dots, w$, определенной в области

$$c \leq y \leq d, \quad l \leq z \leq m, \quad \dots, \quad s \leq w \leq t.$$

Обозначим ее через $F(y, z, \dots, w)$.

Таким образом,

$$F(y, z, \dots, w) = \int_a^b f(x, y, z, \dots, w) dx. \quad (1)$$

Переменные $y, z, \dots, w$, от которых зависит подинтегральная функция и которые при интегрировании рассматриваются как постоянные, называются *параметрами*. Сам интеграл (1) есть функция этих параметров.

Так, например, интеграл $\int_0^1 \sqrt{x+y} dx = \frac{2}{3} [(1+y)^{3/2} - y^{3/2}]$ есть функция одного параметра y , определенная на любом отрезке $c \leq y \leq d$, где $c \geq 0$, или, что то же, для всех $y \geq 0$.

Интеграл $\int_{-2}^1 \frac{dx}{\sqrt{x^2 + y^2 + z^2}} = \ln \frac{1 + \sqrt{1 + y^2 + z^2}}{-2 + \sqrt{4 + y^2 + z^2}}$ есть функция двух параметров y и z , определенная в любом прямоугольнике $\{c \leq y \leq d, l \leq z \leq m\}$, не содержащем начала ($y=0, z=0$)¹⁾; интеграл

$$\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{\sin x (y^2 + z^2 + v^2)}{x} dx$$

есть функция трех параметров y, z, v , определенная для всех значений y, z и v . Этот интеграл в элементарных функциях не берется, если $y^2 + z^2 + v^2 \neq 0$.

Поставим своей задачей изучение некоторых свойств функции $F(y, z, \dots, w)$ по известным свойствам функции $f(x, y, z, \dots, w)$. Для простоты мы будем рассматривать интегралы, зависящие от одного параметра. Полученные результаты могут быть перенесены на случай любого числа параметров.

Теорема 1. Если функция $f(x, y)$ непрерывна в прямоугольнике $\{a \leq x \leq b, c \leq y \leq d\}$, то функция $F(y) = \int_a^\beta f(x, y) dx$, где $a \leq \alpha \leq b$, $a \leq \beta \leq b$, непрерывна на отрезке $c \leq y \leq d$.

Доказательство. Пусть $\alpha < \beta$. Рассмотрим любое значение $y, c \leq y \leq d$. Дадим ему приращение Δy так, чтобы значение $y + \Delta y$ не вышло за пределы отрезка $[c, d]$. При этом функция $F(y)$ получит приращение ΔF :

$$\Delta F = F(y + \Delta y) - F(y) = \int_\alpha^\beta [f(x, y + \Delta y) - f(x, y)] dx.$$

Пусть ε — любое положительное число. Так как функция $f(x, y)$ непрерывна в замкнутой области $\{a \leq x \leq b, c \leq y \leq d\}$, то она равномерно непрерывна в ней, а значит, существует число $\delta, \delta > 0$, такое, что

$$|f(x_1, y_1) - f(x_2, y_2)| < \frac{\varepsilon}{\beta - \alpha}$$

для любых двух точек (x_1, y_1) и (x_2, y_2) области, находящихся на расстоянии, меньшем δ .

Выбирая $|\Delta y| < \delta$, будем иметь:

$$|\Delta F| \leq \int_\alpha^\beta |f(x, y + \Delta y) - f(x, y)| dx \leq \varepsilon,$$

откуда и следует, что функция $F(y)$ непрерывна при любом значении $y, c \leq y \leq d$.

Случай $\alpha > \beta$ сводится к рассмотренному простой перестановкой пределов интегрирования. Для $\alpha = \beta$ утверждение очевидно.

Из доказанной теоремы как следствие вытекает теорема о возможности предельного перехода по параметру под знаком интеграла.

Теорема 2. Если функция $f(x, y)$ непрерывна в прямоугольнике $\{a \leq x \leq b, c \leq y \leq d\}$, то

$$\lim_{y \rightarrow y_0} \int_\alpha^\beta f(x, y) dx = \int_\alpha^\beta \lim_{y \rightarrow y_0} f(x, y) dx,$$

где $c \leq y_0 \leq d, a \leq \alpha \leq b, a \leq \beta \leq b$.

¹⁾ Или, что то же, определенная для всех значений y и z , кроме $y=z=0$.

Доказательство. Действительно, так как (по теореме 1) функция $F(y)$ непрерывна при $y = y_0$, имеем:

$$\begin{aligned} \lim_{y \rightarrow y_0} \int_{\alpha}^{\beta} f(x, y) dx &= \lim_{y \rightarrow y_0} F(y) = F(y_0) = \int_{\alpha}^{\beta} f(x, y_0) dx = \\ &= \int_{\alpha}^{\beta} \lim_{y \rightarrow y_0} f(x, y) dx. \end{aligned}$$

Рассмотрим теперь вопрос о дифференцировании интеграла по параметру.

Теорема 3. Если функции $f(x, y)$ и $f'_y(x, y)$ непрерывны в прямоугольнике $\{a \leq x \leq b, c \leq y \leq d\}$, то функция $F(y) = \int_{\alpha}^{\beta} f(x, y) dx$, где $a \leq \alpha \leq b$, $c \leq \beta \leq d$, дифференцируема на отрезке $[c, d]$, причем

$$F'(y) = \int_{\alpha}^{\beta} f'_y(x, y) dx.$$

Доказательство. Пусть $\alpha < \beta$. Для любого значения y , $c \leq y \leq d$, и приращения Δy (не выводящего за пределы отрезка $[c, d]$), получаем:

$$\Delta F = F(y + \Delta y) - F(y) = \int_{\alpha}^{\beta} [f(x, y + \Delta y) - f(x, y)] dx,$$

или, после применения теоремы Лагранжа,

$$\Delta F = \int_{\alpha}^{\beta} f'_y(x, y + \theta \Delta y) \Delta y dx, \text{ где } \theta = \theta(x, y, \Delta x, \Delta y), 0 < \theta < 1.$$

Разделив обе части равенства на Δy , будем иметь:

$$\frac{\Delta F}{\Delta y} = \int_{\alpha}^{\beta} f'_y(x, y + \theta \Delta y) dx, \text{ где } 0 < \theta < 1.$$

Покажем, что при выполнении условий теоремы

$$\lim_{\Delta y \rightarrow 0} \int_{\alpha}^{\beta} f'_y(x, y + \theta \Delta y) dx = \int_{\alpha}^{\beta} f'_y(x, y) dx.$$

Для этого достаточно показать, что для любого числа $\varepsilon > 0$ существует число $\delta > 0$, такое, что для всех Δy , $|\Delta y| < \delta$, выполнено неравенство:

$$\left| \int_{\alpha}^{\beta} f'_y(x, y + \theta \Delta y) dx - \int_{\alpha}^{\beta} f'_y(x, y) dx \right| < \varepsilon.$$

Пусть $\varepsilon > 0$ — любое число. Так как функция $f'_y(x, y)$ непрерывна в замкнутой области $\{a \leq x \leq b, c \leq y \leq d\}$, то она равномерно непрерывна в ней, а значит, существует число $\delta > 0$, такое, что для любых двух точек этой области (x_1, y_1) и (x_2, y_2) , находящихся на расстоянии, меньшем δ , выполнено неравенство

$$|f'_y(x_1, y_1) - f'_y(x_2, y_2)| < \frac{\varepsilon}{\beta - \alpha}.$$

Выбирая $|\Delta y| < \delta$, будем иметь:

$$\begin{aligned} &\left| \int_{\alpha}^{\beta} f'_y(x, y + \theta \Delta y) dx - \int_{\alpha}^{\beta} f'_y(x, y) dx \right| = \\ &= \left| \int_{\alpha}^{\beta} [f'_y(x, y + \theta \Delta y) - f'_y(x, y)] dx \right| \leq \int_{\alpha}^{\beta} |f'_y(x, y + \theta \Delta y) - f'_y(x, y)| dx < \varepsilon. \end{aligned}$$

Отсюда следует, что для любого значения y , $c \leq y \leq d$, существует производная $F'(y)$, причем

$$F'(y) = \int_a^b f'_y(x, y) dx.$$

Пусть, например, требуется вычислить производную функции $F(y) = \int_0^1 \ln(x^2 + y^2) dx$ при любом значении y , $y > 0$.

Так как условия теоремы 3 выполнены, то при вычислении производной можно дифференцировать по y под знаком интеграла:

$$F'(y) = \int_0^1 \frac{2y}{x^2 + y^2} dx = 2y \cdot \frac{1}{y} \operatorname{arctg} \frac{x}{y} \Big|_{x=0}^{x=1} = 2 \operatorname{arctg} \frac{1}{y}.$$

Дифференцирование интегралов по параметрам иногда применяется для вычисления определенных интегралов.

Пример. Вычислим интеграл

$$\int_0^{\frac{\pi}{2}} \ln(a^2 - \sin^2 \theta) d\theta \quad (a > 1).$$

Здесь θ — переменная интегрирования, a — параметр,

$$\int_0^{\frac{\pi}{2}} \ln(a^2 - \sin^2 \theta) d\theta = F(a).$$

Так как условия теоремы 3 выполнены, для любого $a > 1$ имеем:

$$F'(a) = \int_0^{\frac{\pi}{2}} \frac{2ad\theta}{a^2 - \sin^2 \theta} = \frac{\pi}{\sqrt{a^2 - 1}},$$

откуда, интегрируя по a , найдем:

$$F(a) = \int \frac{\pi da}{\sqrt{a^2 - 1}} = \pi \ln(a + \sqrt{a^2 - 1}) + C.$$

Для определения постоянной C представим интеграл $F(a)$ в виде:

$$F(a) = \int_0^{\frac{\pi}{2}} \left[\ln a^2 + \ln \left(1 - \frac{\sin^2 \theta}{a^2} \right) \right] d\theta = \pi \ln a + \int_0^{\frac{\pi}{2}} \ln \left(1 - \frac{\sin^2 \theta}{a^2} \right) d\theta.$$

Тогда

$$C = -\pi \ln \left(1 + \sqrt{1 - \frac{1}{a^2}} \right) + \int_0^{\frac{\pi}{2}} \ln \left(1 - \frac{\sin^2 \theta}{a^2} \right) d\theta,$$

¹⁾ Для нахождения $F'(y)$ в данном примере можно сначала вычислить $F(y)$, затем продифференцировать. Но этот путь ведет к более громоздким выкладкам.

и, переходя к пределу при $a \rightarrow +\infty$, получим¹⁾ $C = -\pi \ln 2$. Таким образом,

$$F(a) = \int_0^{\frac{\pi}{2}} \ln(a^2 - \sin^2 \theta) d\theta = \pi \ln \frac{a + \sqrt{a^2 - 1}}{2}.$$

Рассматривая интегралы, зависящие от параметров, мы предполагали пределы α и β постоянными. Однако в приложениях часто случается, что при различных значениях параметров интегрировать приходится в разных пределах, т. е. α и β сами являются функциями параметров.

Так, например, интеграл $\int_{\alpha(y)}^{\beta(y)} f(x, y) dx$, где $a \leq \alpha(y) \leq b$, $a \leq \beta(y) \leq b$, есть функция параметра y :

$$\int_{\alpha(y)}^{\beta(y)} f(x, y) dx = F(y).$$

Эта функция при определенных условиях, наложенных на функции $f(x, y)$, $\alpha(y)$ и $\beta(y)$, будет непрерывна и дифференцируема на отрезке $[c, d]$.

Теорема 1'. Если функция $f(x, y)$ непрерывна в прямоугольнике $\{a \leq x \leq b, c \leq y \leq d\}$, а функции $\alpha(y)$ и $\beta(y)$ непрерывны на отрезке $[c, d]$, причем $a \leq \alpha(y) \leq b$, $a \leq \beta(y) \leq b$, то функция $F(y) = \int_{\alpha(y)}^{\beta(y)} f(x, y) dx$ непрерывна на отрезке $[c, d]$.

Доказательства этой теоремы мы не рассматриваем, предоставляя его читателю.

Теорема 3'. Если функции $f(x, y)$ и $f'_y(x, y)$ непрерывны в прямоугольнике $\{a \leq x \leq b, c \leq y \leq d\}$, а функции $\alpha(y)$ и $\beta(y)$ дифференцируемы на отрезке $[c, d]$, причем $a \leq \alpha(y) \leq b$, $a \leq \beta(y) \leq b$, то функция

$$F(y) = \int_{\alpha(y)}^{\beta(y)} f(x, y) dx$$

дифференцируема на отрезке $[c, d]$, причем

$$F'(y) = \int_{\alpha(y)}^{\beta(y)} f'_y(x, y) dx + f(\beta(y), y) \beta'(y) - f(\alpha(y), y) \alpha'(y).$$

Доказательство. Будем рассматривать $F(y)$ как сложную функцию переменной y :

$$F(y) = \Phi(y, \beta, \alpha) = \int_{\alpha}^{\beta} f(x, y) dx, \text{ где } \alpha = \alpha(y), \beta = \beta(y).$$

¹⁾ $\left| \ln \left(1 - \frac{\sin^2 \theta}{a^2} \right) \right| \leq \left| \ln \left(1 - \frac{1}{a^2} \right) \right|$, значит,

$$\left| \int_0^{\frac{\pi}{2}} \ln \left(1 - \frac{\sin^2 \theta}{a^2} \right) d\theta \right| \leq \int_0^{\frac{\pi}{2}} \left| \ln \left(1 - \frac{1}{a^2} \right) \right| d\theta = \left| \ln \left(1 - \frac{1}{a^2} \right) \right| \frac{\pi}{2}.$$

Так как $\lim_{a \rightarrow +\infty} \left| \ln \left(1 - \frac{1}{a^2} \right) \right| \cdot \frac{\pi}{2} = 0$, то и $\lim_{a \rightarrow +\infty} \int_0^{\frac{\pi}{2}} \ln \left(1 - \frac{\sin^2 \theta}{a^2} \right) d\theta = 0$.

Функция $\Phi(y, \beta, \alpha)$ имеет непрерывные частные производные по α , β и y , т. е. дифференцируема по y , причем

$$F'(y) = \frac{\partial \Phi(y, \beta, \alpha)}{\partial y} = \frac{\partial \Phi}{\partial y} + \frac{\partial \Phi}{\partial \beta} \beta'(y) + \frac{\partial \Phi}{\partial \alpha} \alpha'(y)^1).$$

При вычислении $\frac{\partial \Phi}{\partial y}$ переменные α и β считаются постоянными, а значит, по теореме 3

$$\frac{\partial \Phi}{\partial y} = \int_{\alpha}^{\beta} f'_y(x, y) dx.$$

Далее, применяя теорему о дифференцировании определенного интеграла по его верхнему и нижнему пределам, получаем:

$$\frac{\partial \Phi}{\partial \beta} = \frac{\partial}{\partial \beta} \int_{\alpha}^{\beta} f(x, y) dx = f(\beta, y),$$

$$\frac{\partial \Phi}{\partial \alpha} = \frac{\partial}{\partial \alpha} \int_{\alpha}^{\beta} f(x, y) dx = -f(\alpha, y).$$

Таким образом,

$$F'(y) = \int_{\alpha}^{\beta} f'_y(x, y) dx + f(\beta, y) \beta'(y) - f(\alpha, y) \alpha'(y),$$

где $\beta = \beta(y)$, $\alpha = \alpha(y)$.

Примеры.

$$1. \text{ Пусть } F(y) = \int_{y^2}^{3y^2+1} \frac{e^{xy}}{x} dx, \quad y \neq 0.$$

Тогда

$$\begin{aligned} F'(y) &= \int_{y^2}^{3y^2+1} e^{xy} dx + \frac{e^{(3y^2+1)y}}{3y^2+1} 6y - \frac{e^{y^3}}{y^2} 2y = \\ &= \frac{e^{xy}}{y} \Big|_{x=y^2}^{x=3y^2+1} + \frac{e^{(3y^2+1)y}}{3y^2+1} 6y - \frac{2e^{y^3}}{y} = e^{(3y^2+1)y} \left[\frac{1}{y} + \frac{6y}{3y^2+1} \right] - \frac{2e^{y^3}}{y}. \end{aligned}$$

2. Доказать, что интеграл $\frac{1}{\omega} \int_0^x f(t) \sin \omega(x-t) dt$, где $f(t)$ — непрерывная функция, а $\omega = \text{const}$, $\omega \neq 0$, есть решение уравнения $y'' + \omega^2 y = f(x)$. Интеграл $\frac{1}{\omega} \int_0^x f(t) \sin \omega(x-t) dt$ является функцией параметра x , обозначим ее $y(x)$. Дифференцируя этот интеграл по параметру x последовательно два раза, получим:

$$\begin{aligned} y'(x) &= \int_0^x f(t) \cos \omega(x-t) dt + \frac{1}{\omega} f(x) \sin \omega(x-x) = \\ &= \int_0^x f(t) \cos \omega(x-t) dt, \\ y''(x) &= -\omega \int_0^x f(t) \sin \omega(x-t) dt + f(x) \cos \omega(x-x) = \\ &= -\omega \int_0^x f(t) \sin \omega(x-t) dt + f(x) = -\omega^2 y(x) + f(x). \end{aligned}$$

Подставляя $y(x)$ и $y''(x)$ в уравнение $y'' + \omega^2 y = f(x)$, получаем тождество

$$-\omega^2 y(x) + f(x) + \omega^2 y(x) = f(x), \quad f(x) \equiv f(x),$$

откуда следует, что $y(x)$ — решение данного уравнения.

¹⁾ См. теорему о производной сложной функции (гл. VII, § 8).

Раздел 2

КРИВОЛИНЕЙНЫЕ ИНТЕГРАЛЫ ПЕРВОГО И ВТОРОГО РОДА

§ 2. Криволинейные интегралы первого рода (по длине дуги)

Рассмотрим задачу о нахождении массы материальной кривой. Эта задача естественным образом приведет нас к одному из обобщений понятия определенного интеграла — криволинейному интегралу по длине дуги.

Пусть вдоль некоторой кривой непрерывным образом распределена масса¹⁾. *Средней плотностью* какого-либо участка такой кривой называют отношение его массы к длине, а *плотностью распределения массы кривой в данной точке* называют предел средней плотности участка кривой, содержащего эту точку, при стягивании последнего к данной точке.

Поставим своей целью определить массу материальной кривой AB , если известна плотность распределения массы μ в каждой ее точке M : $\mu = \mu(M)$.

Эту задачу можно решить следующим образом. Разобьем дугу AB точками $M_1, M_2, \dots, M_{n-1}$ на n дуг. Обозначим, кроме того, для удобства дальнейших записей, точку A через M_0 , а точку B через M_n ²⁾.

На каждой дуге $M_{i-1} M_i$ выберем произвольную точку N_i и вычислим в этой точке плотность распределения массы кривой:

$$\mu_i = \mu(N_i).$$

Предполагая, что плотность во всех точках дуги $M_{i-1} M_i$ постоянна и равна ее значению в точке N_i , найдем приближенное значение массы m дуги $M_{i-1} M_i$:

$$m_i \approx \mu(N_i) \cdot \Delta s_i,$$

где Δs_i — длина дуги $M_{i-1} M_i$.

Так как масса m всей дуги AB равна сумме масс ее частей:

$$m = \sum_{i=1}^n m_i,$$

то

$$m \approx \sum_{i=1}^n \mu(N_i) \Delta s_i.$$

Полученная формула будет приближенной, так как плотность распределения массы на каждой дуге $M_{i-1} M_i$ в действительности, вообще говоря, не постоянна. Естественно ожидать, что чем меньше будут длины дуг деления, тем точнее будет это приближенное равенство. За массу кривой AB принимают предел полученного приближенного значения при стремлении к нулю всех длин дуг деления:

$$m = \lim_{n \rightarrow \infty} \sum_{i=1}^n \mu(N_i) \Delta s_i.$$

К изучению пределов сумм указанного вида сводится большое количество задач механики и математики, в связи с чем представляется це-

¹⁾ Провода, тросы, цепи и пр. часто рассматриваются в расчетах как материальные кривые (их толщиной пренебрегают по сравнению с длиной).

²⁾ Точки $M_0, M_1, \dots, M_n$ считаем для определенности перенумерованными в направлении от A к B .

лесообразным исследовать эти суммы, отвлекаясь от их конкретного содержания.

Рассмотрим в пространстве XYZ непрерывную спрямляемую кривую AB , в точках которой определена произвольная функция $f(x, y, z)$.

Разобьем кривую AB точками $M_1, \dots, M_{n-1}$, занумерованными для определенности в направлении от A к B на n частей (рис. 351). Точку A обозначим через M_0 , точку B через M_n ¹⁾. На каждой дуге $M_{i-1}M_i$ выберем произвольную точку $N_i(\xi_i, \eta_i, \zeta_i)$ и умножим значение функции $f(x, y, z)$ в этой точке N_i на длину Δs_i дуги $M_{i-1}M_i$:

$$f(\xi_i, \eta_i, \zeta_i) \Delta s_i.$$

Сложив все такие произведения, мы получим сумму:

$$\sum_{i=1}^n f(\xi_i, \eta_i, \zeta_i) \Delta s_i. \quad (1)$$

Ее называют *интегральной суммой* для функции $f(x, y, z)$, заданной на дуге AB . Интегральная сумма зависит от способа разбиения дуги AB на части и от выбора точек N_i на этих частях.

Назовем *шагом разбиения* кривой AB на части наибольшую из длин

дуг деления и обозначим ее через λ . Рассмотрим процесс, при котором число n дуг деления кривой AB будет неограниченно возрастать, а длины этих дуг будут стремиться к нулю. Этот процесс можно характеризовать словами: «шаг разбиения λ кривой AB стремится к нулю».

Если при стремлении к нулю шага разбиения λ существует предел интегральных сумм, то этот предел называют *криволинейным интегралом от функции $f(x, y, z)$ по длине дуги AB* и обозначают символом:

$$\int_{AB} f(x, y, z) ds^2).$$

Криволинейный интеграл по длине дуги часто называют криволинейным интегралом *первого рода* в отличие от криволинейных интегралов второго рода (по координатам), с которыми мы познакомимся в § 3.

Ниже будет показано, что если функция $f(x, y, z)$ непрерывна на спрямляемой непрерывной кривой AB ²⁾, то указанный предел существует, т. е. существует интеграл

$$\int_{AB} f(x, y, z) ds.$$

¹⁾ Можно перенумеровать эти точки и в обратном направлении.

²⁾ Точное определение предела интегральных сумм аналогично соответствующему определению для определенного интеграла. А именно, число I называется пределом интегральных сумм (1), если для любого числа $\epsilon > 0$ существует число $\delta > 0$, такое, что для всякой интегральной суммы с шагом разбиения, меньшим δ , выполнено неравенство

$$\left| I - \sum_{i=1}^n f(\xi_i, \eta_i, \zeta_i) \Delta s_i \right| < \epsilon.$$

Этот предел, как следует из определения, не зависит от способа разбиения кривой AB на части и от выбора точек N_i на этих частях.

³⁾ При этом функция $f(x, y, z) = f(M)$ называется *непрерывной в точке P* кривой AB , если для любой последовательности точек $P_1, P_2, \dots, P_n, \dots$ кривой AB , сходящейся к точке P , выполнено равенство: $\lim_{n \rightarrow \infty} f(P_n) = f(P)$. Функция, непрерывная в каждой точке кривой AB , называется *непрерывной на кривой AB* .

Сформулируем некоторые свойства криволинейного интеграла по длине дуги, непосредственно вытекающие из определения этого интеграла. При этом каждый из упомянутых ниже интегралов будем предполагать существующим.

1. Значение криволинейного интеграла по длине дуги не зависит от направления кривой AB .

2. Постоянный множитель можно выносить за знак криволинейного интеграла.

3. Криволинейный интеграл от суммы двух функций равен сумме интегралов от этих функций.

4. Если кривая AB точкой C разбита на две части AC и CB , то

$$\int_{AB} f(x, y, z) ds = \int_{AC} f(x, y, z) ds + \int_{CB} f(x, y, z) ds.$$

5. Если в точках кривой AB

$$f_1(x, y, z) \leq f_2(x, y, z),$$

то

$$\int_{AB} f_1(x, y, z) ds \leq \int_{AB} f_2(x, y, z) ds.$$

6.

$$\left| \int_{AB} f(x, y, z) ds \right| \leq \int_{AB} |f(x, y, z)| ds.$$

7. Если $f(x, y, z) \equiv 1$, то

$$\int_{AB} ds = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \Delta s_i = S,$$

где S — длина кривой AB .

8. Теорема о среднем. Если функция $f(x, y, z)$ непрерывна на кривой AB , то на этой кривой существует точка (ξ, η, ζ) , такая, что:

$$\int_{AB} f(x, y, z) ds = f(\xi, \eta, \zeta) \cdot S,$$

где S — длина кривой AB .

Возвращаясь к рассмотренной выше механической задаче, заключаем, что масса m материальной кривой AB вычисляется по формуле

$$m = \int_{AB} \mu(x, y, z) ds,$$

где $\mu(x, y, z)$ — плотность распределения массы.

Выясним теперь, как связан криволинейный интеграл по длине дуги с обыкновенным определенным интегралом.

Пусть AB — непрерывная спрямляемая кривая, а $f(x, y, z)$ — функция, непрерывная на этой кривой.

Установим на кривой AB определенное направление, например от A к B . Тогда положение любой точки $M(x, y, z)$ кривой AB можно характеризовать длиной s дуги AM :

$$x = x(s), \quad y = y(s), \quad z = z(s).$$

В частности, точке A соответствует $s=0$, а точке B соответствует $s=S$, где S — длина кривой AB . При этом на уравнения $x=x(s)$, $y=y(s)$, $z=z(s)$, где s меняется в пределах от $s=0$ до $s=S$, можно смотреть как на параметрические уравнения кривой AB .

Функция $f(x, y, z) = f(x(s), y(s), z(s))$ становится сложной непрерывной функцией параметра s , $0 \leq s \leq S$. Обозначая через σ_i и σ_i зна-

чения параметра s , соответствующие точкам M_i и N_i , $s_{i-1} \leq \sigma_i \leq s_i$, получим:

$$\sum_{i=1}^n f(\xi_i, \eta_i, \zeta_i) \Delta s_i = \sum_{i=1}^n f(x(\sigma_i), y(\sigma_i), z(\sigma_i)) \Delta s_i,$$

где $\Delta s_i = s_i - s_{i-1}$ — длина дуги $M_{i-1}M_i$.

В правой части равенства стоит интегральная сумма для обыкновенного определенного интеграла от непрерывной функции $f(x(s), y(s), z(s))$ на отрезке $0 \leq s \leq S$. Переходя к пределу при стремлении к нулю шага разбиения λ кривой AB , получаем:

$$\int_{AB} f(x, y, z) ds = \int_0^S f(x(s), y(s), z(s)) ds. \quad (2)$$

Заметим, что эти рассуждения не только дают выражение (2) криволинейного интеграла по длине дуги через обыкновенный интеграл, но и доказывают существование криволинейного интеграла от функции $f(x, y, z)$, непрерывной на рассматриваемой кривой AB .

Пусть теперь кривая AB задана параметрическими уравнениями $x = x(t)$, $y = y(t)$, $z = z(t)$, $\alpha \leq t \leq \beta$, где $x(t)$, $y(t)$, $z(t)$ — непрерывно дифференцируемые функции любого параметра t (не обязательно длины дуги s), причем точке A соответствует $t = \alpha$, а точке B $t = \beta$. Установим на кривой AB направление от A к B . Тогда длина дуги AM [$M(x, y, z)$ — текущая точка кривой AB] будет функцией параметра t :

$$s = s(t) = \int_{\alpha}^t \sqrt{x'^2(t) + y'^2(t) + z'^2(t)} dt.$$

Производя замену переменной в правой части формулы (2), будем иметь:

$$\int_{AB} f(x, y, z) ds = \int_{\alpha}^{\beta} f(x(t), y(t), z(t)) \sqrt{x'^2(t) + y'^2(t) + z'^2(t)} dt. \quad (3)$$

Если точке A соответствует большее значение параметра ($t = \beta$), чем точке B ($t = \alpha$), то, устанавливая на кривой AB направление от B к A и производя аналогичные выкладки, снова приходим к формуле (3), так как интеграл по длине дуги от направления кривой не зависит.

Таким образом, для вычисления криволинейного интеграла по длине дуги надо, используя параметрические уравнения кривой, выразить подинтегральную функцию через параметр t , заменить ds дифференциалом дуги как функции параметра t и проинтегрировать полученное выражение по t в пределах от меньшего до большего значения параметра t для данной кривой.

Пусть, например, требуется вычислить интеграл $\int_{AB} (x^2 + y^2 + z^2) ds$ по одному витку винтовой линии: $x = \cos t$, $y = \sin t$, $z = t$, $0 \leq t \leq 2\pi$ (рис. 352).

По формуле (3) получаем:

$$\begin{aligned} \int_{AB} (x^2 + y^2 + z^2) ds &= \int_0^{2\pi} (\cos^2 t + \sin^2 t + t^2) \sqrt{(-\sin t)^2 + (\cos t)^2 + 1} dt = \\ &= \sqrt{2} \int_0^{2\pi} (1 + t^2) dt = 2\sqrt{2} \pi \left(1 + \frac{4\pi^2}{3}\right). \end{aligned}$$

Если AB — плоская кривая, заданная уравнением $y = \varphi(x)$, $a \leq x \leq b$, где $\varphi(x)$ — непрерывно дифференцируемая функция, то, принимая x за параметр, как частный случай формулы (3) будем иметь формулу

$$\int_{AB} f(x, y) ds = \int_a^b f(x, \varphi(x)) \sqrt{1 + \varphi'^2(x)} dx. \quad (4)$$

Пример. Вычислить интеграл $\int_{AB} x^2 ds$, где AB — часть логарифмической кривой $y = \ln x$ от $x=1$ до $x=2$ (рис. 353).

По формуле (4)

$$\int_{AB} x^2 ds = \int_1^2 x^2 \sqrt{1 + \frac{1}{x^2}} dx = \frac{1}{3} (1 + x^2)^{\frac{3}{2}} \Big|_{x=1}^{x=2} = \frac{1}{3} (5\sqrt{5} - 2\sqrt{2}).$$

Криволинейный интеграл по длине дуги имеет разнообразные приложения в механике и математике. С помощью этого интеграла можно

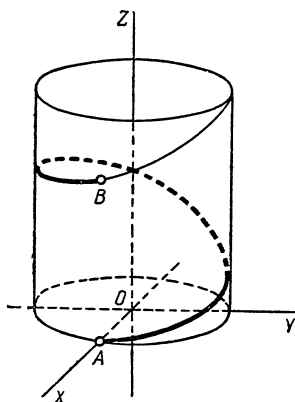


Рис. 352.

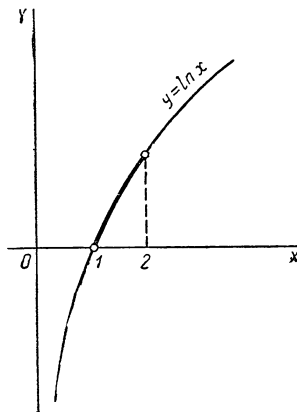


Рис. 353.

вычислить массу материальной кривой, ее статические моменты относительно координатных осей и плоскостей, моменты инерции относительно координатных осей, найти координаты центра тяжести такой кривой, решить задачу о притяжении материальной точки материальной кривой и т. д.

§ 3. Криволинейные интегралы второго рода

Прежде чем дать определение криволинейного интеграла второго рода, рассмотрим приводящую к нему задачу о вычислении работы переменной силы при перемещении материальной точки вдоль некоторой кривой.

Будем говорить, что в области V пространства XYZ задано силовое поле с напряженностью F , если в каждой точке M области определена сила F , действующая на единицу массы, помещенную в эту точку. При этом величина и направление силы F предполагаются зависящими только от положения точки M : $F = F(M)$. Тогда проекции P, Q, R силы F на оси координат будут функциями координат x, y, z точки M :

$$P = P(x, y, z), \quad Q = Q(x, y, z), \quad R = R(x, y, z).$$

Вычислим работу, совершаемую силой F по перемещению единицы массы вдоль непрерывной спрямляемой кривой AB в направлении от A к B (короче, вычислим работу поля вдоль кривой AB).

Из механики известно, что если перемещение прямолинейно, а действующая сила F постоянная по величине и направлению, то затраченная ею работа A равна скалярному произведению силы F на вектор смещения l :

$$A = (F, l).$$

В случае непрямолинейного пути перемещения и переменной силы F определение выполненной работы можно произвести следующим образом.

Разобьем кривую AB в направлении от A к B точками $M_0, M_1, \dots, M_n$, где точка M_0 совпадает с A , M_n — с B , на n дуг.

Пусть x_i, y_i, z_i — координаты точки M_i . На каждой дуге $M_{i-1}M_i$ возьмем произвольную точку $N_i(\xi_i, \eta_i, \zeta_i)$. Предположим, что в точках дуги $M_{i-1}M_i$ действующая сила F постоянна и равна ее значению в точке N_i , и предположим также, что под действием этой силы материальная точка перемещалась не по дуге $M_{i-1}M_i$, а по отрезку $M_{i-1}M_i$ (рис. 354).

Работу постоянной силы $F(N_i)$ вдоль отрезка $M_{i-1}M_i$ можно принять за приближенное значение работы A_i поля вдоль дуги $M_{i-1}M_i$.

Таким образом, $A_i \approx (F(N_i), \overline{M_{i-1}M_i})$ или, так как вектор $\overline{M_{i-1}M_i}$ имеет координаты:

$$\Delta x_i = x_i - x_{i-1}, \Delta y_i = y_i - y_{i-1}, \Delta z_i = z_i - z_{i-1},$$

то

$$A_i \approx P(\xi_i, \eta_i, \zeta_i) \Delta x_i + Q(\xi_i, \eta_i, \zeta_i) \Delta y_i + R(\xi_i, \eta_i, \zeta_i) \Delta z_i.$$

Считая, что работа A поля вдоль дуги AB равна сумме работ вдоль ее частей: $A = \sum_{i=1}^n A_i$, получаем:

$$A \approx \sum_{i=1}^n [P(\xi_i, \eta_i, \zeta_i) \Delta x_i + Q(\xi_i, \eta_i, \zeta_i) \Delta y_i + R(\xi_i, \eta_i, \zeta_i) \Delta z_i].$$

За точное значение работы по перемещению единичной массы вдоль кривой AB из точки A в точку B принимается предел, к которому стремится найденное приближенное значение при стремлении к нулю наибольшей из длин дуг деления λ :

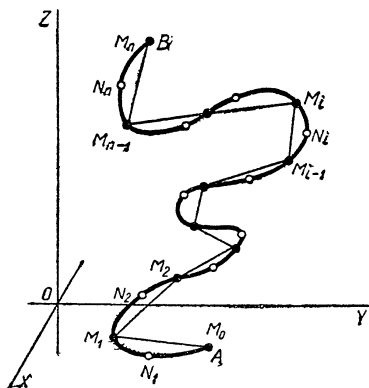


Рис. 354.

$$A = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n [P(\xi_i, \eta_i, \zeta_i) \Delta x_i + Q(\xi_i, \eta_i, \zeta_i) \Delta y_i + R(\xi_i, \eta_i, \zeta_i) \Delta z_i].$$

Будем изучать пределы указанного вида, отвлекаясь от конкретного содержания рассмотренной задачи.

Итак, пусть AB — непрерывная кривая в пространстве XYZ (в частности, на плоскости XOY), а $P(x, y, z)$ — произвольная функция, определенная на этой кривой.

Разобьем дугу AB точками $M_0, M_1, \dots, M_n$ в направлении от A к B на n дуг¹⁾. Через x_i, y_i, z_i обозначим координаты точек $M_i, i = 0, 1, 2, \dots, n$. На каждой дуге $M_{i-1}M_i$ выберем произвольную точку $N_i(\xi_i, \eta_i, \zeta_i)$ и умножим значение функции $P(x, y, z)$ в точке N_i на величину $\Delta x_i = x_i - x_{i-1}$ проекции дуги $M_{i-1}M_i$ на ось OX .

Сумма всех таких произведений:

$$\sum_{i=1}^n P(\xi_i, \eta_i, \zeta_i) \Delta x_i$$

называется *интегральной суммой* для функции $P(x, y, z)$ на дуге AB по переменной x .

Если при стремлении к нулю шага разбиения λ кривой AB такие интегральные суммы имеют предел, то этот предел называют *криво-*

¹⁾ Здесь, как и раньше, точка A обозначена через M_0 , а B — через M_n .

линейным интегралом по переменной x от функции $P(x, y, z)$ по кривой AB в направлении от A к B и обозначают символом

$$\int_{AB} P(x, y, z) dx^1). \quad (1)$$

Итак, $\int_{AB} P(x, y, z) dx = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n P(\xi_i, \eta_i, \zeta_i) \Delta x_i$, если этот предел существует.

Криволинейный интеграл по переменной x отличается от криволинейного интеграла по длине дуги тем, что при составлении интегральной суммы значение функции в точке N_i дуги $M_{i-1}M_i$ умножается не на длину дуги $M_{i-1}M_i$, а на величину Δx_i ее проекции на ось OX ²⁾. Так как, в отличие от длины дуги, величина проекции дуги $M_{i-1}M_i$ на ось OX существенно зависит от направления дуги и меняет знак при перемене направления, заключаем, что криволинейный интеграл по переменной x зависит от направления кривой AB и при перемене направления кривой меняет знак на обратный:

$$\int_{AB} P(x, y, z) dx = - \int_{BA} P(x, y, z) dx$$

(здесь AB и BA обозначают одну и ту же кривую, на которой в первом случае установлено направление от A к B , а во втором случае — противоположное направление, от B к A).

Аналогичным образом для любых функций $Q(x, y, z)$ и $R(x, y, z)$, заданных на кривой AB , можно определить *криволинейные интегралы по этой кривой в направлении от A к B по переменным y и z* :

$$\int_{AB} Q(x, y, z) dy = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n Q(\xi_i, \eta_i, \zeta_i) \Delta y_i, \quad (2)$$

$$\int_{AB} R(x, y, z) dz = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i) \Delta z_i, \quad (3)$$

если эти пределы существуют. Криволинейные интегралы по переменным x, y и z называют *криволинейными интегралами второго рода*.

Сумму криволинейных интегралов

$$\int_{AB} P(x, y, z) dx, \int_{AB} Q(x, y, z) dy, \int_{AB} R(x, y, z) dz$$

также называют криволинейным интегралом второго рода и обозначают:

$$\int_{AB} P(x, y, z) dx + Q(x, y, z) dy + R(x, y, z) dz. \quad (4)$$

Каждый из этих криволинейных интегралов существенно зависит от направления кривой, по которой производится интегрирование, и меняет знак при перемене направления пути интегрирования, т. е.

$$\int_{AB} = - \int_{BA}^3).$$

Рассмотрим свойства криволинейных интегралов второго рода, непосредственно вытекающие из их определения. При этом каждый из встречающихся ниже интегралов будем предполагать существующим.

¹⁾ Точное определение предела интегральных сумм указанного вида предоставляется дать читателю по аналогии с соответствующим определением для определенного интеграла.

²⁾ Величиной проекции направленной дуги $M_{i-1}M_i$ на ось OX называют величину проекции направленного отрезка $\overline{M_{i-1}M_i}$ на эту ось.

³⁾ Подинтегральные выражения мы для краткости опускаем.

1. Постоянный множитель выносится за знак криволинейного интеграла:

$$\int_{AB} kPdx = k \int_{AB} Pdx, \quad \int_{AB} kQdy = k \int_{AB} Qdy, \quad \int_{AB} kRdz = k \int_{AB} Rdz,$$

где k — число.

2. Криволинейный интеграл от суммы двух функций равен сумме соответствующих интегралов от слагаемых:

$$\begin{aligned} \int_{AB} (P_1 + P_2) dx &= \int_{AB} P_1 dx + \int_{AB} P_2 dx, \\ \int_{AB} (Q_1 + Q_2) dy &= \int_{AB} Q_1 dy + \int_{AB} Q_2 dy, \\ \int_{AB} (R_1 + R_2) dz &= \int_{AB} R_1 dz + \int_{AB} R_2 dz. \end{aligned}$$

3. Если кривая AB точкой C разбита на две части, то криволинейный интеграл по всей кривой AB равен сумме интегралов по ее частям:

$$\int_{AB} = \int_{AC} + \int_{CB}.$$

4. Криволинейный интеграл по замкнутой кривой L не зависит от выбора начальной точки, а зависит только от направления обхода кривой.

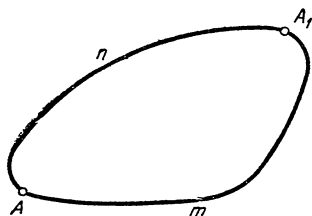


Рис. 355.

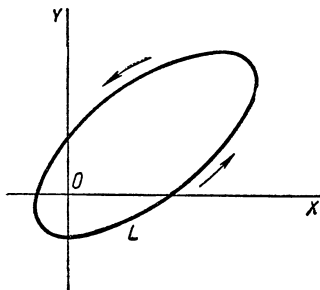


Рис. 356.

Действительно, выбирая точку A за начальную (точка A совпадает с конечной точкой B), получаем (см. рис. 355): $\int_{AmA_1nA} = \int_{AmA_1} + \int_{A_1nA}$, где A_1 — любая точка кривой L , отличная A .

Выбирая теперь точку A_1 за начальную (точка A_1 совпадает с конечной точкой B_1), получаем:

$$\int_{A_1nAmA_1} = \int_{A_1nA} + \int_{AmA_1}.$$

Таким образом,

$$\int_{A_1nAmA_1} = \int_{AmA_1nA}.$$

Криволинейный интеграл по замкнутой кривой L (по контуру L) обычно обозначают символом $\oint_L$. Крючок на знаке интеграла служит для обозначения контура. Направление обхода контура L задается дополнительно.

В случае, если L — замкнутая кривая без точек самопересечения, лежащая в плоскости XOY , условимся называть *положительным направ-*

¹⁾ Свойства 3 и 4 имеют место для криволинейных интегралов вида (1) — (4)

лением обхода контура L такое направление обхода, при котором движение происходит против часовой стрелки, т. е. так, что ограниченная контуром L часть плоскости остается слева (рис. 356). Противоположное направление обхода контура будем называть *отрицательным*. Соответствующие криволинейные интегралы принято обозначать символами:

$$\oint_L \text{ и } \oint_L^-.$$

При отсутствии дополнительных указаний о направлении обхода контура L , лежащего в плоскости XOY , условимся подразумевать положительное направление обхода.

5. Если AB — кривая, лежащая в плоскости, перпендикулярной оси OX , то

$$\int_{AB} P(x, y, z) dx = 0.$$

Действительно, в этом случае в соответствующей интегральной сумме все Δx_i , $i = 1, 2, \dots, n$, равны нулю, а значит,

$$\int_{AB} P(x, y, z) dx = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n P(\xi_i, \eta_i, \zeta_i) \Delta x_i = 0.$$

Аналогично для кривой AB , лежащей в плоскости, перпендикулярной оси OY (OZ), получаем:

$$\int_{AB} Q(x, y, z) dy = 0 \quad \left(\int_{AB} R(x, y, z) dz = 0 \right).$$

Сформулируем теперь **достаточные условия существования криволинейных интегралов второго рода**.

Будем называть кривую AB *гладкой кривой*, если существуют параметрические уравнения этой кривой: $x = x(t)$, $y = y(t)$, $z = z(t)$, $t_1 \leq t \leq t_2$, где $x(t)$, $y(t)$, $z(t)$ — непрерывно дифференцируемые функции на отрезке $[t_1, t_2]$. Непрерывная кривая AB называется *кусочно-гладкой*, если ее можно разбить на конечное число гладких дуг¹⁾.

Теорема. Если кривая AB — кусочно-гладкая, а функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ непрерывны на кривой AB , то криволинейные интегралы

$$\int_{AB} P(x, y, z) dx, \quad \int_{AB} Q(x, y, z) dy, \quad \int_{AB} R(x, y, z) dz,$$

а следовательно, и криволинейный интеграл

$$\int_{AB} P(x, y, z) dx + Q(x, y, z) dy + R(x, y, z) dz,$$

существуют.

Доказательство этой теоремы в данном курсе не рассматривается.

Возвращаясь к рассмотренной выше задаче о работе поля $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ вдоль кривой AB , заключаем, что при определенных условиях²⁾, наложенных на кривую и на функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$, эта работа вычисляется по формуле

$$A = \int_{AB} P(x, y, z) dx + Q(x, y, z) dy + R(x, y, z) dz.$$

Вычисление криволинейных интегралов второго рода производится путем их преобразования в определенные интегралы.

¹⁾ Как следует из определения, всякую кусочно-гладкую кривую можно задать параметрическими уравнениями: $x = x(t)$, $y = y(t)$, $z = z(t)$, $t_1 \leq t \leq t_2$, где функции $x(t)$, $y(t)$, $z(t)$ непрерывны на отрезке $[t_1, t_2]$, а их производные $x'(t)$, $y'(t)$, $z'(t)$ непрерывны всюду, кроме, может быть, конечного числа точек разрыва первого рода.

²⁾ См., например, условия сформулированной выше теоремы.

Пусть кривая AB задана параметрическими уравнениями $x=x(t)$, $y=y(t)$, $z=z(t)$, где $x(t)$, $y(t)$, $z(t)$ — непрерывно дифференцируемые функции t , изменяющегося в пределах от $t=\alpha$ до $t=\beta$ ($\alpha > \beta$ или $\alpha < \beta$). Предположим, что точке A соответствует значение параметра $t=\alpha$, точке B соответствует $t=\beta$, причем изменению t от α к β соответствует движение точки $M(x(t), y(t), z(t))$ по кривой AB в направлении от A к B .

Тогда, если функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ непрерывны на кривой AB , то криволинейные интегралы

$$\int_{AB} P(x, y, z) dx, \quad \int_{AB} Q(x, y, z) dy \quad \text{и} \quad \int_{AB} R(x, y, z) dz,$$

определяемые как пределы соответствующих интегральных сумм, существуют и, следовательно, не зависят от способов построения этих сумм. Воспользовавшись этим, построим интегральные суммы специальным образом.

Разобьем отрезок с концами $t=\alpha$ и $t=\beta$ оси Ot точками $t_0=\alpha$, t_1 , t_2 , ..., $t_n=\beta$ в направлении от α к β на n частей. Обозначим через $M_i(x_i, y_i, z_i)$ точку кривой AB , соответствующую значению параметра t_i :

$$x_i = x(t_i), \quad y_i = y(t_i), \quad z_i = z(t_i).$$

Точки $M_0, M_1, \dots, M_n$ разобьют кривую AB в направлении от A к B на n дуг. По теореме Лагранжа

$$\Delta x_i = x_i - x_{i-1} = x(t_i) - x(t_{i-1}) = x'(\tau_i) \Delta t_i,$$

где $\Delta t_i = t_i - t_{i-1}$, а τ_i — значение параметра t , лежащее между t_{i-1} и t_i .

Пусть $N_i(\xi_i, \eta_i, \zeta_i)$ — точка кривой AB , соответствующая значению параметра $t=\tau_i$: $\xi_i = x(\tau_i)$, $\eta_i = y(\tau_i)$, $\zeta_i = z(\tau_i)$. Эта точка лежит на дуге $M_{i-1}M_i$.

Составим интегральную сумму для интеграла $\int_{AB} P(x, y, z) dx$, используя полученное разбиение кривой AB на части $M_{i-1}M_i$ ($i=1, 2, \dots, n$) и выбранные точки N_i на этих частях. По построению имеем:

$$\sum_{i=1}^n P(\xi_i, \eta_i, \zeta_i) \Delta x_i = \sum_{i=1}^n P(x(\tau_i), y(\tau_i), z(\tau_i)) x'(\tau_i) \Delta t_i. \quad (5)$$

Правая часть равенства представляет собой интегральную сумму для определенного интеграла от непрерывной функции $P(x(t), y(t), z(t)) x'(t)$ по переменному t , изменяющемуся от $t=\alpha$ до $t=\beta$. При стремлении к нулю шага разбиения рассматриваемого промежутка оси Ot , в силу непрерывности функций $x'(t)$, $y'(t)$, $z'(t)$, будет стремиться к нулю шаг разбиения кривой AB . Переходя к пределу в равенстве (5), получим формулу, выражающую криволинейный интеграл по переменной x через определенный интеграл:

$$\int_{AB} P(x, y, z) dx = \int_{\alpha}^{\beta} P(x(t), y(t), z(t)) x'(t) dt. \quad (6)$$

Аналогичным образом выводятся формулы:

$$\int_{AB} Q(x, y, z) dy = \int_{\alpha}^{\beta} Q(x(t), y(t), z(t)) y'(t) dt, \quad (7)$$

$$\int_{AB} R(x, y, z) dz = \int_{\alpha}^{\beta} R(x(t), y(t), z(t)) z'(t) dt. \quad (8)$$

В общем случае получаем:

$$\int_{AB} P dx + Q dy + R dz = \int_{\alpha}^{\beta} [Px'(t) + Qy'(t) + Rz'(t)] dt. \quad (9)$$

Таким образом, для вычисления криволинейного интеграла по переменным x , y и z по кривой AB надо составить параметрические уравнения пути интегрирования, выразить формально под знаком интеграла переменные x , y и z , а также их дифференциалы dx , dy и dz через параметр, и затем вычислить от полученного выражения определенный интеграл в пределах изменения параметра от значения, соответствующего началу пути интегрирования, до значения, соответствующего концу.

Примечание. Выведенные формулы также имеют место, если кривая AB кусочно-гладкая (а не гладкая, как предполагалось при доказательстве), а функции $x(t)$, $y(t)$, $z(t)$ удовлетворяют указанным на стр. 509 условиям (см. сноску 1).

Рассмотрим несколько примеров на вычисление криволинейных интегралов второго рода.

Пример 1. Вычислить интеграл $\int_{AB} x^2 dx - yz dy + z dz$ вдоль отрезка AB , соединяющего точки $A(1, 2, -1)$ и $B(3, 3, 2)$.

Уравнения прямой AB :

$$\frac{x-1}{2} = \frac{y-2}{1} = \frac{z+1}{3}$$

или в параметрической форме: $x = 1 + 2t$, $y = 2 + t$, $z = -1 + 3t$.

Для отрезка AB параметр t изменяется от $t_A = 0$ до $t_B = 1$. Следовательно,

$$\begin{aligned} \int_{AB} x^2 dx - yz dy + z dz &= \int_0^1 [(1+2t)^2 \cdot 2 - (2+t)(-1+3t) + (-1+3t)3] dt = \\ &= \int_0^1 (5t^2 + 12t + 1) dt = \frac{26}{3}. \end{aligned}$$

В данном примере за параметр также можно взять любую из переменных x , y или z .

Например, выбирая за параметр y , запишем уравнения отрезка AB в следующей форме:

$$x = 2y - 3, \quad y = y, \quad z = 3y - 7, \quad y_A = 2 \leq y \leq y_B = 3.$$

Применяя выведенную выше формулу, получим:

$$\begin{aligned} \int_{AB} x^2 dx - yz dy + z dz &= \\ &= \int_2^3 [(2y-3)^2 \cdot 2 - y(3y-7) + (3y-7)3] dy = \frac{26}{3}. \end{aligned}$$

Пример 2. Вычислить интеграл $\oint (x+y+z) dx$ по ломаной $ABCA$ с вершинами $A(1, 0, 0)$, $B(0, 1, 0)$, $C(0, 0, 1)$ (рис. 357).

По определению $\oint_{ABCA} = \int_{AB} + \int_{BC} + \int_{CA}$.

$\int_{BC} (x+y+z) dx = 0$, так как путь интегрирования лежит в плоскости, перпендикулярной оси Ox . Уравнения отрезка AB запишем в виде: $x = x$, $y = 1 - x$, $z = 0$.

Так как $x_A = 1$, $x_B = 0$, имеем $\int_{AB} (x+y+z) dx = \int_1^0 [x + (1-x)] dx = -1$.

Уравнения отрезка CA запишем в виде: $x = x$, $y = 0$, $z = 1 - x$. Так как $x_C = 0$, $x_A = 1$, имеем:

$$\int_{CA} (x+y+z) dx = \int_0^1 [x + (1-x)] dx = 1.$$

В результате получаем:

$$\oint_{ABCA} (x+y+z) dx = -1 + 0 + 1 = 0.$$

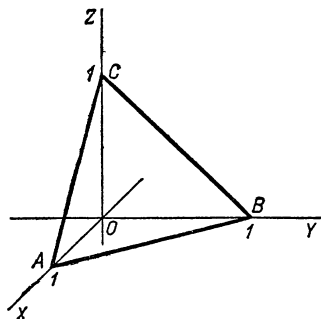


Рис. 357.

В случае, если AB — плоская кривая, заданная уравнением $y=f(x)$, $f(x)$ — непрерывно дифференцируемая функция x , и функции $P(x, y)$ и $Q(x, y)$ непрерывны на этой кривой, то, принимая за параметр x , из формул (6) — (9), как следствие, получаем формулы:

$$\int_{AB} P(x, y) dx = \int_{x_A}^{x_B} P(x, f(x)) dx, \quad (6')$$

$$\int_{AB} Q(x, y) dy = \int_{x_A}^{x_B} Q(x, f(x)) f'(x) dx, \quad (7')$$

$$\begin{aligned} \int_{AB} P(x, y) dx + Q(x, y) dy &= \\ &= \int_{x_A}^{x_B} [P(x, f(x)) + Q(x, f(x)) f'(x)] dx. \end{aligned} \quad (9')$$

Здесь x_A и x_B — значения переменной x соответственно для точек A и B .

Пример 1. Вычислить интеграл: $\int_{AB} xy dx - (x-y) dy$ по дуге AB параболы $y=x^2$, если $x_A=-1$, $x_B=2$. Принимая за параметр x , получим:

$$\begin{aligned} \int_{AB} xy dx - (x-y) dy &= \int_{-1}^2 [x^3 - (x-x^2) \cdot 2x] dx = \\ &= \left[\frac{3}{4} x^4 - \frac{2}{3} x^3 \right]_{-1}^2 = 15 \frac{11}{12}. \end{aligned}$$

Пример 2. Вычислить интеграл $\int_{AOB} y dx$, если AOB — дуга параболы $y^2=x$,

причем $A(1, -1)$, $B(1, 1)$ (рис. 358). Учитывая свойства криволинейных интегралов, получим:

$$\int_{AOB} y dx = \int_{AO} y dx + \int_{OB} y dx.$$

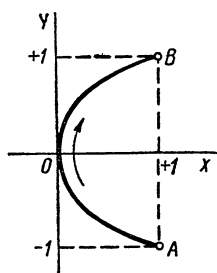


Рис. 358.

Так как дуга AO имеет уравнение $y=-\sqrt{x}$, $0 \leq x \leq 1$, причем $x_A=1$, $x_O=0$, дуга OB имеет уравнение $y=\sqrt{x}$, $0 \leq x \leq 1$, $x_O=0$, $x_B=1$, то

$$\int_{AOB} y dx = \int_1^0 -\sqrt{x} dx + \int_0^1 \sqrt{x} dx = 2 \int_0^1 \sqrt{x} dx = \frac{4}{3}.$$

Примечание. В данном примере проще рассматривать в качестве параметра y , видоизменив соответственно формулу (7):

$$\int_{AOB} y dx = \int_{-1}^1 y \cdot 2y dy = \frac{2}{3} y^3 \Big|_{-1}^1 = \frac{4}{3}.$$

Рассмотрим коротко вопрос о связи криволинейных интегралов первого и второго рода. Пусть AB — гладкая кривая, а функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ непрерывны на ней. Установим на кривой AB направление от A к B и выберем в качестве параметра длину s дуги AM , где $M(x, y, z)$ — текущая точка кривой. При этом $s_A=0$, $s_B=S$ (S — длина дуги AB) и параметрические уравнения кривой запишутся в виде:

$$x=x(s), \quad y=y(s), \quad z=z(s), \quad 0 \leq s \leq S.$$

Если обозначить через α , β , γ углы, составляемые с осями координат касательной к кривой AB , направленной в сторону роста параметра s , то, как можно показать,

$$\cos \alpha = x'(s), \quad \cos \beta = y'(s), \quad \cos \gamma = z'(s)^1).$$

¹⁾ Вывод этих формул для плоской кривой приведен в § 25 гл. IV.

Учитывая эти соотношения и формулы, выражающие криволинейные интегралы первого и второго рода через определенные интегралы, получим:

$$\begin{aligned}\int_{AB} P(x, y, z) dx &= \int_0^S P(x(s), y(s), z(s)) x'(s) ds = \int_{AB} P(x, y, z) \cos \alpha ds, \\ \int_{AB} Q(x, y, z) dy &= \int_0^S Q(x(s), y(s), z(s)) y'(s) ds = \int_{AB} Q(x, y, z) \cos \beta ds, \\ \int_{AB} R(x, y, z) dz &= \int_0^S R(x(s), y(s), z(s)) z'(s) ds = \int_{AB} R(x, y, z) \cos \gamma ds.\end{aligned}$$

Таким образом, криволинейные интегралы второго рода выражаются через криволинейные интегралы по длине дуги, т. е. первого рода.

Для общего криволинейного интеграла второго рода имеем формулу:

$$\int_{AB} P dx + Q dy + R dz = \int_{AB} (P \cos \alpha + Q \cos \beta + R \cos \gamma) ds^1).$$

§ 4. Условия независимости криволинейного интеграла (второго рода) от формы пути интегрирования

Пусть функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ непрерывны в некоторой области D пространства XYZ . Будем рассматривать в этой области только кусочно-гладкие кривые.

Возьмем в области D две произвольные точки A и B . Их можно соединять различными кривыми, лежащими в области D . По каждой из этих кривых интеграл $\int_{AB} P dx + Q dy + R dz$ имеет, вообще говоря, свое значение.

Вычислим, например, интеграл $\int_{AB} x dx + xy dy + y dz$ по отрезку прямой AB , соединяющему точки $A(1, 0, 0)$ и $B(1, 1, 1)$ (рис. 359). Уравнения прямой AB : $\frac{x-1}{0} = \frac{y-0}{1} = \frac{z-0}{1}$, т. е. $x=1$, $y=z$.

Выбирая за параметр y , имеем:

$$\int_{AB} x dx + xy dy + y dz = \int_0^1 2y dy = 1.$$

Вычислим теперь тот же интеграл по дуге параболы AB , заданной уравнениями $x=1$, $z=y^2$ (рис. 359). Выбирая за параметр y ($x=1$, $y=y$, $z=y^2$, $y_A=0$, $y_B=1$), получим:

$$\int_{AB} x dx + xy dy + y dz = \int_0^1 (y + 2y^3) dy = \frac{7}{6}.$$

Этот пример показывает, что значения интеграла $\int_{AB} x dx + xy dy + y dz$ зависят от формы кривой, по которой производится интегрирование.

Можно, однако, привести примеры криволинейных интегралов, значения которых по любым кривым, соединяющим данные точки A и B ,

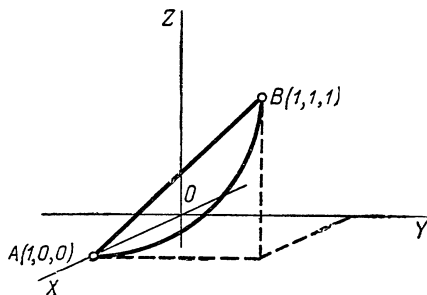


Рис. 359.

¹⁾ Аналогичные равенства имеют место и для кусочно-гладкой кривой AB .

будут одни и те же. Такими, как можно показать, являются интегралы:

$$\int_{AB} (yz + 1) dx + (xz - 2) dy + xy dz, \quad \int_{AB} x dx + y dy + z dz \quad \text{и др.}$$

Будем говорить, что интеграл $\int_{AB} P dx + Q dy + R dz$ в области D не зависит от формы пути интегрирования, если его значения по всевозможным кусочно-гладким кривым, лежащим в данной области и имеющим общее начало и общий конец, одинаковы. В этом случае при записи интеграла достаточно указывать лишь начальную и конечную точки пути интегрирования, в связи с чем приняты обозначения:

$$\int_A^B P dx + Q dy + R dz \quad \text{или} \quad \int_{(x_A, y_A, z_A)}^{(x_B, y_B, z_B)} P dx + Q dy + R dz.$$

Независимость криволинейного интеграла $\int_{AB} P dx + Q dy + R dz$ в области D от формы пути интегрирования равносильна равенству нулю этого интеграла вдоль всякой замкнутой кривой L , лежащей в области D .

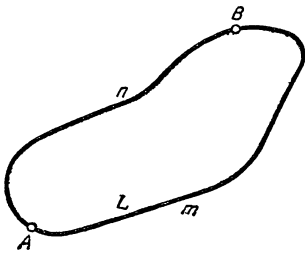


Рис. 360.

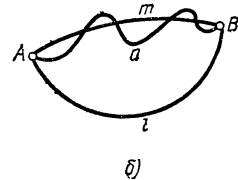
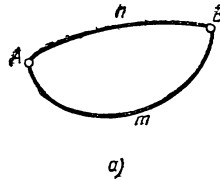


Рис. 361.

Действительно, для любой замкнутой кривой L , лежащей в этой области, имеем $\oint_L = \int_{AmB} + \int_{BnA}$, где A и B — две любые точки на кривой L (см. рис. 360).

Если интеграл $\int_{AB} P dx + Q dy + R dz$ в области D не зависит от формы пути интегрирования, то

$$\int_{AmB} = \int_{AnB} = - \int_{BnA}, \quad \text{т. е.} \quad \int_{AmB} + \int_{BnA} = \oint_L = 0.$$

Наоборот, пусть интеграл $\oint_L P dx + Q dy + R dz$ равен нулю по любой замкнутой кривой L в области D . Возьмем любые две точки A и B в области D и соединим их любыми двумя кривыми AmB и AnB . Если эти кривые пересекаются только в точках A и B , то, рассматривая замкнутую кривую $AmBnA$ (рис. 361, а), получим:

$$\oint_{AmBnA} = 0 = \int_{AmB} + \int_{BnA} = \int_{AmB} - \int_{AnB}, \quad \text{т. е.} \quad \int_{AmB} = \int_{AnB}.$$

Если же кривые AmB и AnB пересекаются в точках, отличных от точек A и B , рассмотрим третью кривую AlB , пересекающуюся с кривыми AmB и AnB только в точках A и B (рис. 361, б). По доказанному

$$\int_{AmB} = \int_{AlB}, \quad \int_{AnB} = \int_{AlB}, \quad \text{т. е.} \quad \int_{AmB} = \int_{AnB}$$

и в этом случае.

Рассмотрим теперь условия, при которых криволинейный интеграл $\int_{AB} P dx + Q dy + R dz$ в области D не зависит от формы пути интегрирования.

Теорема. Для того чтобы криволинейный интеграл $\int_{AB} P dx + Q dy + R dz$, где $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ — функции, непрерывные в области D , в этой области не зависел от формы пути, необходимо и достаточно, чтобы выражение $P dx + Q dy + R dz$ было полным дифференциалом некоторой функции (в области D).

Доказательство. Достаточность. Предположим, что существует функция $U(x, y, z)$, такая, что в области D всюду

$$P dx + Q dy + R dz = dU(x, y, z).$$

Покажем, что в этом случае в области D интеграл $\int_{AB} P dx + Q dy + R dz$ не зависит от формы пути интегрирования. Пусть A и B — две любые фиксированные точки области D , AB — любая кривая, соединяющая эти точки. Если $x = x(t)$, $y = y(t)$, $z = z(t)$ — уравнения этой кривой, то

$$\begin{aligned} \int_{AB} P dx + Q dy + R dz &= \\ &= \int_{t_A}^{t_B} dU(x(t), y(t), z(t)) = \\ &= U(x(t_B), y(t_B), z(t_B)) - \\ &\quad - U(x(t_A), y(t_A), z(t_A)) = \\ &= U(x_B, y_B, z_B) - U(x_A, y_A, z_A), \end{aligned}$$

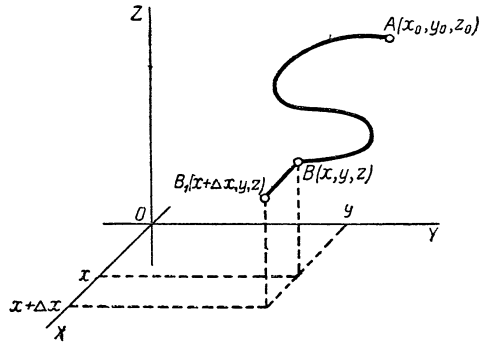


Рис. 362.

т. е. интеграл зависит лишь от положения точек A и B и не зависит от формы кривой AB .

Необходимость. Предположим теперь, что интеграл $\int_{AB} P dx + Q dy + R dz$ в области D не зависит от формы пути интегрирования. Фиксируем в области D точку $A(x_0, y_0, z_0)$, а точку $B(x, y, z)$ рассматриваем как переменную.

Криволинейный интеграл $\int_{(x_0, y_0, z_0)}^{(x, y, z)} P dx + Q dy + R dz$ будет зависеть от положения точки $B(x, y, z)$, т. е. будет функцией переменных x, y, z , определенной в области D . Обозначим эту функцию через $U(x, y, z)$:

$$\int_{(x_0, y_0, z_0)}^{(x, y, z)} P dx + Q dy + R dz = U(x, y, z),$$

и покажем, что подинтегральное выражение

$$P dx + Q dy + R dz$$

равно дифференциалу функции $U(x, y, z)$ в области D .

Пусть $B(x, y, z)$ — любая фиксированная точка области D . Дадим x приращение Δx . Тогда

$$U(x + \Delta x, y, z) = \int_{(x_0, y_0, z_0)}^{(x + \Delta x, y, z)} P dx + Q dy + R dz.$$

Так как величина криволинейного интеграла не зависит от формы пути, будем рассматривать интеграл $\int_{(x_0, y_0, z_0)}^{(x, y, z)}$ по какой-нибудь кривой AB , а

интеграл $\int_{(x_0, y_0, z_0)}^{(x+\Delta x, y, z)}$ — по кривой AB_1 , состоящей из выбранной кривой AB и отрезка прямой BB_1 (B_1 — точка с координатами $x+\Delta x, y, z$) (рис. 362). При этом

$$\begin{aligned}\Delta_x U &= U(x+\Delta x, y, z) - U(x, y, z) = \int_{(x_0, y_0, z_0)}^{(x+\Delta x, y, z)} P dx + Q dy + R dz - \\ &- \int_{(x_0, y_0, z_0)}^{(x, y, z)} P dx + Q dy + R dz = \int_{(x, y, z)}^{(x+\Delta x, y, z)} P dx + Q dy + R dz = \\ &= \int_{(x, y, z)}^{(x+\Delta x, y, z)} P(x, y, z) dx,\end{aligned}$$

так как отрезок BB_1 лежит в плоскости, перпендикулярной осям OY и OZ . Выражая последний интеграл через определенный и применяя теорему о среднем к определенному интегралу, получим:

$$\Delta_x U = \int_x^{x+\Delta x} P(t, y, z) dt = P(x+\theta\Delta x, y, z)\Delta x,$$

где $0 \leq \theta \leq 1$. Деля обе части этого равенства на Δx , в силу непрерывности функции $P(x, y, z)$ будем иметь:

$$\frac{\partial U(x, y, z)}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta_x U}{\Delta x} = \lim_{\Delta x \rightarrow 0} P(x+\theta\Delta x, y, z) = P(x, y, z).$$

Аналогично доказывается, что

$$\frac{\partial U(x, y, z)}{\partial y} = Q(x, y, z) \text{ и } \frac{\partial U(x, y, z)}{\partial z} = R(x, y, z).$$

Так как функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ непрерывны в области D , то частные производные $\frac{\partial U}{\partial x}$, $\frac{\partial U}{\partial y}$, $\frac{\partial U}{\partial z}$ функции $U(x, y, z)$ непрерывны в той же области. Отсюда следует, что в каждой точке области D функция $U(x, y, z)$ дифференцируема и что

$$dU(x, y, z) = P dx + Q dy + R dz.$$

Выясним теперь, какие условия надо наложить на функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$, чтобы выражение $P dx + Q dy + R dz$ представляло собой полный дифференциал некоторой функции.

Ответ на этот вопрос дается следующей теоремой.

Теорема. Если функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ в области $D \{a \leq x \leq b, c \leq y \leq d, l \leq z \leq k\}$ ¹⁾ имеют непрерывные частные производные первого порядка, то, для того чтобы выражение $P dx + Q dy + R dz$ в этой области было полным дифференциалом некоторой функции, необходимым и достаточным является выполнение условий:

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y} \quad (1)$$

в каждой точке области D .

Доказательство. Необходимость. Предположим, что всюду в области D

$$P dx + Q dy + R dz = dU,$$

¹⁾ В частности, этот параллелепипед может быть неограниченным.

где U — некоторая функция переменных x, y, z . Тогда $\frac{\partial U}{\partial x} = P, \frac{\partial U}{\partial y} = Q, \frac{\partial U}{\partial z} = R$ в каждой точке области, а значит,

$$\begin{aligned}\frac{\partial^2 U}{\partial x \partial y} &= \frac{\partial P}{\partial y}, & \frac{\partial^2 U}{\partial y \partial x} &= \frac{\partial Q}{\partial x}, \\ \frac{\partial^2 U}{\partial x \partial z} &= \frac{\partial P}{\partial z}, & \frac{\partial^2 U}{\partial z \partial x} &= \frac{\partial R}{\partial x}, \\ \frac{\partial^2 U}{\partial y \partial z} &= \frac{\partial Q}{\partial z}, & \frac{\partial^2 U}{\partial z \partial y} &= \frac{\partial R}{\partial y}.\end{aligned}$$

Так как частные производные функций P, Q, R , записанные в правых частях этих равенств, в области D непрерывны, то по теореме о независимости смешанных производных от последовательности дифференцирований заключаем, что

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}$$

в области D .

Как следует из доказательства, это утверждение остается в силе, если указанную в формулировке теоремы область D заменить любой областью пространства XYZ .

Аналогичным образом устанавливается, что если в некоторой области n -мерного пространства функции

$$P(x, y, z, \dots, w), \quad Q(x, y, z, \dots, w), \dots, M(x, y, z, \dots, w)$$

непрерывны вместе со своими частными производными первого порядка, то, для того чтобы выражение $P dx + Q dy + \dots + M dw$ было полным дифференциалом некоторой функции $U(x, y, z, \dots, w)$, необходимо выполнение в этой области C_n^2 условий:

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \dots, \frac{\partial P}{\partial w} = \frac{\partial M}{\partial x}, \dots,$$

представляющих собой равенства смешанных производных второго порядка от функции $U(x, y, z, \dots, w)$, отличающихся лишь последовательностью дифференцирований.

Достаточность. Предположим теперь, что равенства (1) выполнены в каждой точке области $D \{a \leq x \leq b, c \leq y \leq d, l \leq z \leq k\}$. Покажем, что существует функция $U(x, y, z)$, такая, что выражение $P dx + Q dy + R dz$ является ее полным дифференциалом, т. е. такая, что $\frac{\partial U}{\partial x} = P, \frac{\partial U}{\partial y} = Q, \frac{\partial U}{\partial z} = R$ в каждой точке области D . Пусть (x_0, y_0, z_0) — некоторая точка области D . Интегрируя первое из соотношений (1) по x , найдем функцию $U(x, y, z)$, удовлетворяющую условию $\frac{\partial U}{\partial x} = P(x, y, z)$:

$$U(x, y, z) = \int_{x_0}^x P(x, y, z) dx + \varphi(y, z). \quad (2)$$

Здесь $\varphi(y, z)$ — произвольная функция переменных y и z . В частности, ее можно считать дифференцируемой. Вычислим частные производные $\frac{\partial U}{\partial y}$ и $\frac{\partial U}{\partial z}$ от функции $U(x, y, z)$, определенной равенством (2):

$$\begin{aligned}\frac{\partial U}{\partial y} &= \frac{\partial}{\partial y} \int_{x_0}^x P(x, y, z) dx + \varphi'_y(y, z), \\ \frac{\partial U}{\partial z} &= \frac{\partial}{\partial z} \int_{x_0}^x P(x, y, z) dx + \varphi'_z(y, z).\end{aligned}$$

Применяя теорему о дифференцировании интеграла по параметру, получаем:

$$\frac{\partial U}{\partial y} = \int_{x_0}^x \frac{\partial P(x, y, z)}{\partial y} dx + \varphi'_y(y, z), \quad \frac{\partial U}{\partial z} = \int_{x_0}^x \frac{\partial P(x, y, z)}{\partial z} dx + \varphi'_z(y, z),$$

или, учитывая первые два из равенств (1):

$$\frac{\partial U}{\partial y} = \int_{x_0}^x \frac{\partial Q(x, y, z)}{\partial x} dx + \varphi'_y(y, z) = Q(x, y, z) - Q(x_0, y, z) + \varphi'_y(y, z),$$

$$\frac{\partial U}{\partial z} = \int_{x_0}^x \frac{\partial R(x, y, z)}{\partial x} dx + \varphi'_z(y, z) = R(x, y, z) - R(x_0, y, z) + \varphi'_z(y, z).$$

Для того чтобы функция $U(x, y, z)$ удовлетворяла условиям

$$\frac{\partial U}{\partial y} = Q(x, y, z) \quad \text{и} \quad \frac{\partial U}{\partial z} = R(x, y, z),$$

функцию $\varphi(y, z)$ следует выбрать так, чтобы

$$\varphi'_y(y, z) = Q(x_0, y, z), \quad \varphi'_z(y, z) = R(x_0, y, z). \quad (3)$$

Интегрируя по y первое из этих соотношений, находим:

$$\varphi(y, z) = \int_{y_0}^y Q(x_0, y, z) dy + \psi(z),$$

где $\psi(z)$ — произвольная дифференцируемая функция z . Отсюда

$$\begin{aligned} \varphi'_z(y, z) &= \frac{\partial}{\partial z} \int_{y_0}^y Q(x_0, y, z) dy + \psi'(z) = \\ &= \int_{y_0}^y \frac{\partial Q(x_0, y, z)}{\partial z} dy + \psi'(z) = \int_{y_0}^y \frac{\partial R(x_0, y, z)}{\partial y} dy + \psi'(z) = \\ &= R(x_0, y, z) - R(x_0, y_0, z) + \psi'(z) \end{aligned}$$

[мы применяли здесь снова правило дифференцирования интеграла по параметру и использовали последнее из условий (1) теоремы].

Функция $\varphi(y, z)$ будет удовлетворять соотношению $\varphi'_z(y, z) = R(x_0, y, z)$,

если $\psi'(z) = R(x_0, y_0, z)$, т. е. если $\psi(z) = \int_{z_0}^z R(x_0, y_0, z) dz + C$, где C — произвольная постоянная.

Таким образом, найдена функция

$$\begin{aligned} U(x, y, z) &= \int_{x_0}^x P(x, y, z) dx + \int_{y_0}^y Q(x_0, y, z) dy + \\ &+ \int_{z_0}^z R(x_0, y_0, z) dz + C, \end{aligned}$$

где C — любое число, удовлетворяющая условиям

$$\frac{\partial U}{\partial x} = P(x, y, z), \quad \frac{\partial U}{\partial y} = Q(x, y, z), \quad \frac{\partial U}{\partial z} = R(x, y, z).$$

Так как функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ непрерывны в области D , то и частные производные $\frac{\partial U}{\partial x}$, $\frac{\partial U}{\partial y}$ и $\frac{\partial U}{\partial z}$ функции $U(x, y, z)$ непрерывны. Отсюда следует, что функция $U(x, y, z)$ дифференцируема и

$$dU = P dx + Q dy + R dz$$

в каждой точке области D .

Примечание. С помощью дополнительных рассуждений можно показать, что теорема имеет место во всякой области, в которой две любые кривые с общим началом и общим концом можно непрерывно деформировать одну в другую, не выходя за пределы области.

Предложенный выше метод доказательства теоремы (см. доказательство достаточности) обычно используется на практике для отыскания функции по ее полному дифференциалу или, как говорят, при интегрировании полного дифференциала.

Например, выражение $(2xy^3 - 2z)dx + (3x^2y^2 + 1)dy + (3 - 2x)dz$ является полным дифференциалом некоторой функции $U(x, y, z)$ во всем пространстве XYZ , так как

$$\begin{aligned}\frac{\partial}{\partial y}(2xy^3 - 2z) &= 6xy^2 = \frac{\partial}{\partial x}(3x^2y^2 + 1), \\ \frac{\partial}{\partial z}(2xy^3 - 2z) &= -2 = \frac{\partial}{\partial x}(3 - 2x), \\ \frac{\partial}{\partial z}(3x^2y^2 + 1) &= 0 = \frac{\partial}{\partial y}(3 - 2x)\end{aligned}$$

во всех точках пространства.

Для определения функции $U(x, y, z)$ имеем соотношение

$$(2xy^3 - 2z)dx + (3x^2y^2 + 1)dy + (3 - 2x)dz = dU,$$

т. е. три соотношения вида:

$$\frac{\partial U}{\partial x} = 2xy^3 - 2z, \quad \frac{\partial U}{\partial y} = 3x^2y^2 + 1, \quad \frac{\partial U}{\partial z} = 3 - 2x. \quad (4)$$

Интегрируя первое из них по x , находим: $U = x^2y^3 - 2xz + \varphi(y, z)$, где $\varphi(y, z)$ — любая функция y и z . При этом $\frac{\partial U}{\partial y} = 3x^2y^2 + \varphi'_y(y, z)$, $\frac{\partial U}{\partial z} = -2x + \varphi'_z(y, z)$. Сравнивая полученные частные производные с соотношениями (4), видим, что $\varphi'_y(y, z) = 1$, $\varphi'_z(y, z) = 3$, откуда следует, что $\varphi(y, z) = y + \psi(z)$, где $\psi(z)$ — любая функция z ,

$$\varphi'_z(y, z) = \psi'(z), \quad \psi'(z) = 3, \quad \psi(z) = 3z + C,$$

где C — произвольная постоянная.

Таким образом, $U = x^2y^3 - 2xz + y + 3z + C$, где C — произвольная постоянная, — искомая функция.

Частными случаями доказанных теорем являются следующие результаты.

1. Для того чтобы в данной области плоскости XOY криволинейный интеграл $\int_{AB} P dx + Q dy$, где $P(x, y)$ и $Q(x, y)$ — непрерывные функции, не зависел от формы пути интегрирования, необходимо и достаточно, чтобы в этой области выражение $P dx + Q dy$ было полным дифференциалом некоторой функции $U(x, y)$.

2. Для того чтобы в данной области¹⁾ плоскости XOY выражение $P dx + Q dy$, где функции $P(x, y)$, $Q(x, y)$ и их частные производные первого порядка непрерывны в области D , было полным дифференциалом, необходимо и достаточно, чтобы в этой области выполнялось условие $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.

Рассмотренные в этом разделе теоремы позволяют легко и просто решать вопрос о зависимости или независимости криволинейного интеграла от формы пути интегрирования.

Так, например, интеграл $\int_{AB} e^y dx - z dy + xyz dz$ в любой области зависит от формы пути, так как подинтегральное выражение не является полным дифференциалом. Здесь $P(x, y, z) = e^y$, $Q(x, y, z) = -z$, $R(x, y, z) = xyz$ и, например, $\frac{\partial P}{\partial y} \neq \frac{\partial Q}{\partial x}$.

¹⁾ Область должна удовлетворять условиям, сформулированным в примечании к доказанной теореме.

Интеграл $\int_{AB} (yz + 1) dx + (xz - 2) dy + xy dz$ не зависит от формы пути интегрирования, а зависит только от положения точек A и B , так как

$$\frac{\partial (yz + 1)}{\partial y} = z = \frac{\partial (xz - 2)}{\partial x}, \quad \frac{\partial (yz + 1)}{\partial z} = y = \frac{\partial (xy)}{\partial x}, \quad \frac{\partial (xz - 2)}{\partial z} = x = \frac{\partial (xy)}{\partial y}.$$

Для вычисления этого интеграла при конкретно заданных точках A и B можно произвольно выбрать путь интегрирования (например, взять отрезок прямой AB или ломаную со звеньями, параллельными осям координат, идущую от точки A к точке B), а можно отыскать предварительно функцию $U(x, y, z)$, для которой $dU = (yz + 1) dx + (xz - 2) dy + xy dz$, а затем вычислить интеграл по формуле

$$\int_{AB} P dx + Q dy + R dz = \int_{AB} dU = U(x_B, y_B, z_B) - U(x_A, y_A, z_A).$$

Вернемся к рассмотренной выше механической задаче. Как было показано, работа силового поля $\mathbf{F} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$ вдоль кривой AB (при определенных условиях) выражается криволинейным интегралом по кривой AB :

$$A = \int_{AB} P dx + Q dy + R dz.$$

Силовое поле называется *потенциальным* в данной области, если в этой области работа поля не зависит от формы пути или, что то же, если *циркуляция* (работа) поля по любому замкнутому пути равна нулю.

По доказанному, для того чтобы поле в данной области было потенциальным, необходимо и достаточно, чтобы в этой области выражение $P dx + Q dy + R dz$ было полным дифференциалом некоторой функции $U(x, y, z)$.

Таким образом, равенства $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, $\frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}$, $\frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}$ представляют собой *условия потенциальности* векторного поля.

Функция $U(x, y, z)$ называется *силовой*, или *потенциальной*, функцией поля.

В потенциальном векторном поле работа поля вдоль любого пути AB вычисляется по формуле

$$A = U(x_B, y_B, z_B) - U(x_A, y_A, z_A),$$

т. е. равна приращению силовой функции поля при переходе из точки A в точку B .

Примером потенциального поля, как легко проверить, может служить любое центрированное поле $\mathbf{F} = f(r)\mathbf{r}$, где $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ — радиус-вектор точки $M(x, y, z)$, $r = \sqrt{x^2 + y^2 + z^2}$ — модуль этого радиуса-вектора, а $f(r)$ — любая дифференцируемая функция r .

Раздел 3 ДВОЙНЫЕ ИНТЕГРАЛЫ

§ 5. Задача об объеме цилиндрического тела

Двойные интегралы представляют собой одно из возможных обобщений понятия определенного интеграла на случай функции двух переменных. К введению этого понятия естественным образом приводит решение задачи об объеме цилиндрического тела.

Рассмотрим в пространстве XYZ тело V , ограниченное сверху поверхностью S , заданной уравнением $z = f(x, y)$, где $f(x, y)$ — непрерывная

функция, сбоку — цилиндрической поверхностью с образующими, параллельными оси OZ , снизу — областью σ плоскости XOY (σ — проекция поверхности S на плоскость XOY). Будем называть тело V цилиндрическим телом с основанием σ (рис. 363). Требуется вычислить объем этого цилиндрического тела¹⁾. Для решения поставленной задачи разобьем область σ на n частей: $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ и рассмотрим цилиндрические столбики с основаниями $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$, ограниченные сверху кусками поверхности S . В своей совокупности они составляют тело V .

Для определения объема ΔV_i столбика с основанием $\Delta\sigma_i$ возьмем в области $\Delta\sigma_i$ произвольно точку $P_i(\xi_i, \eta_i)$ и построим цилиндр с основанием $\Delta\sigma_i$ и высотой $h_i = f(\xi_i, \eta_i)$ (рис. 364). Объем этого цилиндра, равный произведению $f(\xi_i, \eta_i) \Delta\sigma_i$ ²⁾, можно принять за приближенное значение объема ΔV_i :

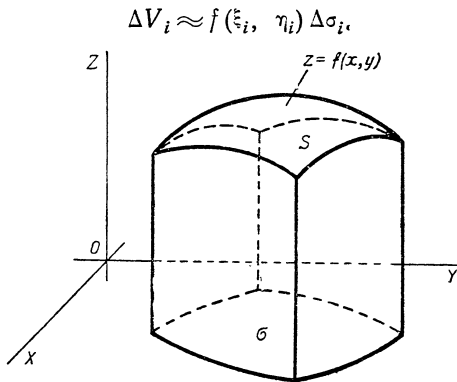


Рис. 363.

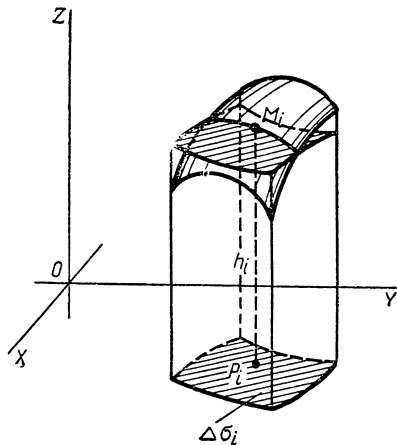


Рис. 364.

Производя указанные вычисления для каждой области деления и складывая результаты, получим приближенное значение объема данного цилиндрического тела:

$$V \approx \sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i.$$

Естественно ожидать, что точное значение объема цилиндрического тела равно пределу, к которому стремится найденное приближенное значение объема при неограниченном измельчении областей деления и, следовательно, неограниченном увеличении числа n областей деления:

$$V = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i.$$

К вычислению пределов сумм аналогичного вида приводит большое количество математических и прикладных задач (некоторые из них ниже будут рассмотрены), в связи с чем представляет интерес их более подробное изучение.

¹⁾ Определение объема тела см. в гл. VI, § 11.

²⁾ Условимся в дальнейшем области и их площади, тела и их объемы обозначать соответственно одинаковыми буквами.

³⁾ В § 8 будет показано, что рассматриваемое цилиндрическое тело кубировуемо и что его объем действительно определяется полученным равенством.

§ 6. Определение двойного интеграла

Пусть в замкнутой квадратуемой¹⁾ области σ плоскости XOY задана функция $f(x, y)$ ²⁾. Разобьем область σ произвольно на n частей $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$, так чтобы области $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ не имели общих внутренних точек (рис. 365). В каждой замкнутой области $\Delta\sigma_i$ (внутри или на границе) выберем произвольную точку $P_i(\xi_i, \eta_i)$ и умножим значение функции $f(x, y)$ в этой точке на площадь $\Delta\sigma_i$. Сложив все такие произведения, получим сумму:

$$\sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i, \quad (1)$$

которая называется *интегральной суммой* для функции $f(x, y)$ в области σ . Интегральная сумма зависит от способа разбиения области σ на части $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ и от выбора точек P_i на этих частях, т. е. для функции $f(x, y)$ в области σ можно составить бесчисленное множество интегральных сумм.

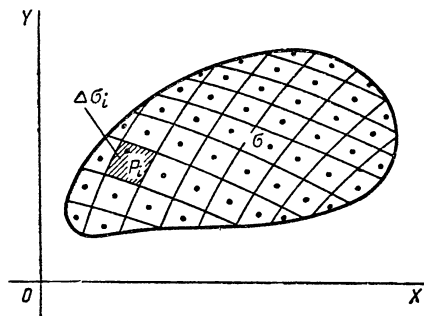


Рис. 365.

Прежде чем переходить к операции предельного перехода, введем понятия диаметра области и шага разбиения области. *Диаметром* замкнутой области (двухмерного или трехмерного пространства) называется наибольшее из расстояний между двумя точками границы этой области. *Шагом разбиения* области на конечное число частей называется наибольший из диаметров областей деления.

Обозначим через λ шаг разбиения области σ на части $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$. Процесс, состоящий в неограниченном измельчении областей деления и соответственно в неограниченном увеличении числа n областей деления, удобно характеризовать словами: «шаг разбиения области σ стремится к нулю».

Если при стремлении к нулю шага разбиения λ области σ интегральные суммы (1) имеют предел, то этот предел называют *двойным интегралом* от функции $f(x, y)$ по области σ и обозначают символами:

$$\iint_{\sigma} f(x, y) d\sigma \quad \text{или} \quad \iint_{\sigma} f(x, y) dx dy^3).$$

Здесь $f(x, y)$ — *подинтегральная функция*, σ — *область интегрирования*, x и y — *переменные интегрирования*, $d\sigma (dx dy)$ — *элемент площади*.

Таким образом, по определению

$$\iint_{\sigma} f(x, y) d\sigma = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i,$$

если этот предел существует.

¹⁾ В дальнейшем, если не оговорено противное, будут рассматриваться только *квадрируемые* (имеющие конечную площадь) области.

²⁾ Заметим, что никаких предположений о непрерывности функции или о ее знаке не делается.

³⁾ При этом число I называется *пределом* интегральных сумм (2) при стремлении λ к нулю, если для любого числа $\varepsilon > 0$ существует число $\delta > 0$, такое, что для всякой интегральной суммы с шагом разбиения, меньшим δ , выполнено неравенство

$$\left| I - \sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i \right| < \varepsilon.$$

Этот предел, как следует из определения, не зависит от способа разбиения области σ на части и от выбора точек на этих частях.

Функция $f(x, y)$, для которой существует двойной интеграл

$$\iint_{\sigma} f(x, y) d\sigma,$$

называется *интегрируемой* в области σ .

§ 7. Условия существования двойного интеграла и его свойства

Естественно возникает вопрос, для всяких ли функций, и если нет, то для каких, существует по области σ двойной интеграл.

Очевидно, что интегрируемая в области σ функция должна быть ограничена в замкнутой области σ , так как в противном случае за счет выбора точек P_i интегральную сумму можно было бы сделать сколько угодно большой по абсолютной величине, а значит, не существовало бы предела интегральных сумм при стремлении λ к нулю¹⁾.

Сформулируем теперь **достаточные условия существования двойного интеграла**.

Теорема 1. Если функция $f(x, y)$ непрерывна в замкнутой области σ , то двойной интеграл $\iint_{\sigma} f(x, y) d\sigma$ существует.

Доказательство этой теоремы можно провести аналогично доказательству соответствующей теоремы для определенного интеграла.

Теорема 2. Если функция $f(x, y)$ ограничена в замкнутой области σ и непрерывна в ней всюду, кроме конечного числа кусочно-гладких линий, то двойной интеграл $\iint_{\sigma} f(x, y) d\sigma$ существует.

Доказательство этой теоремы в данном курсе не приводится.

Предполагая, что все встречающиеся ниже функции удовлетворяют условиям теорем (1 или 2) существования двойного интеграла в рассматриваемых областях, перечислим простейшие свойства двойного интеграла. Доказательства этих свойств непосредственно следуют из определения двойного интеграла. Читатель легко докажет свойства 1—7 по аналогии с соответствующими доказательствами для определенного интеграла.

1. Двойной интеграл $\iint_{\sigma} f(x, y) dx dy$ не зависит от обозначения переменных интегрирования.

2. Постоянный множитель k можно выносить за знак двойного интеграла:

$$\iint_{\sigma} kf(x, y) d\sigma = k \iint_{\sigma} f(x, y) d\sigma.$$

3. Двойной интеграл от суммы двух функций равен сумме двойных интегралов от слагаемых:

$$\iint_{\sigma} [f_1(x, y) + f_2(x, y)] d\sigma = \iint_{\sigma} f_1(x, y) d\sigma + \iint_{\sigma} f_2(x, y) d\sigma^3).$$

4. Если область σ разбита на две области σ_1 и σ_2 , то

$$\iint_{\sigma} f(x, y) d\sigma = \iint_{\sigma_1} f(x, y) d\sigma + \iint_{\sigma_2} f(x, y) d\sigma^3).$$

¹⁾ См. доказательство соответствующего утверждения для определенного интеграла (гл. VI, разд. 1, § 3).

²⁾ Это свойство справедливо для любого фиксированного числа слагаемых.

³⁾ Свойство 7 обобщается на случай, когда область σ разбита на любое число m областей $\sigma_1, \sigma_2, \dots, \sigma_m$.

5. Если всюду в области σ

$$f(x, y) \geq 0,$$

то

$$\iint_{\sigma} f(x, y) d\sigma \geq 0.$$

6. Если всюду в области σ

$$f_1(x, y) \leq f_2(x, y),$$

то

$$\iint_{\sigma} f_1(x, y) d\sigma \leq \iint_{\sigma} f_2(x, y) d\sigma.$$

7. Если функция $f(x, y)$ задана в области σ , то

$$\left| \iint_{\sigma} f(x, y) d\sigma \right| \leq \iint_{\sigma} |f(x, y)| d\sigma.$$

8.

$$\iint_{\sigma} d\sigma = \sigma.$$

Действительно, так как в данном случае $f(x, y, z) = 1$ в области σ , то для любого разбиения области σ на части $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ получим:

$$\iint_{\sigma} d\sigma = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \Delta\sigma_i = \lim_{\lambda \rightarrow 0} \sigma = \sigma.$$

Свойство 8 позволяет использовать двойные интегралы для нахождения площадей плоских фигур.

§ 8. Геометрический смысл двойного интеграла

Вернемся к задаче об объеме цилиндрического тела. Так как рассматриваемая там функция $f(x, y)$ непрерывна в замкнутой области σ , то

$$\lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i = \iint_{\sigma} f(x, y) d\sigma,$$

каковы бы ни были точки $P_i(\xi_i, \eta_i)$ в замкнутых областях $\Delta\sigma_i$. Выбирая за $P_i(\xi_i, \eta_i)$ точки, в которых функция $f(x, y)$ в областях $\Delta\sigma_i$ достигает своих наименьших (m_i) или соответственно наибольших (M_i) значений, будем, в частности, иметь:

$$\iint_{\sigma} f(x, y) d\sigma = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n m_i \Delta\sigma_i = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n M_i \Delta\sigma_i.$$

Слагаемые $m_i \Delta\sigma_i$ и $M_i \Delta\sigma_i$ равны объемам цилиндров с основанием $\Delta\sigma_i$, вписанных и описанных вокруг ограниченного сверху куском поверхности S цилиндрического столбика с тем же основанием $\Delta\sigma_i$ (рис. 366).

Суммы $\sum_{i=1}^n m_i \Delta\sigma_i$ и $\sum_{i=1}^n M_i \Delta\sigma_i$ равны соответственно объемам «вписанных» и «описанных» вокруг данного цилиндрического тела V тел, состоящих из n прямых цилиндров. Так как при $\lambda \rightarrow 0$ эти объемы имеют общий предел, данное тело кубируемо и его объем V равен двойному интегралу:

$$V = \iint_{\sigma} f(x, y) d\sigma.$$

Отсюда следует *геометрический смысл двойного интеграла*: двойной интеграл $\iint_{\sigma} f(x, y) d\sigma$ от неотрицательной, непрерывной в замкнутой области σ функции $f(x, y)$ равен объему цилиндрического тела с основанием σ в плоскости XOY , ограниченного сверху поверхностью $z=f(x, y)$.

В частности, если $f(x, y) \equiv 1$ в области σ , то цилиндрическое тело представляет собой цилиндр с основанием σ высоты $h=1$ (рис. 367). Его объем численно равен площади основания σ . Таким образом, мы получаем уже известную формулу для вычисления площади области σ с помощью двойного интеграла $\sigma = \iint_{\sigma} d\sigma$ или, что

то же, $\sigma = \iint_{\sigma} dx dy$.

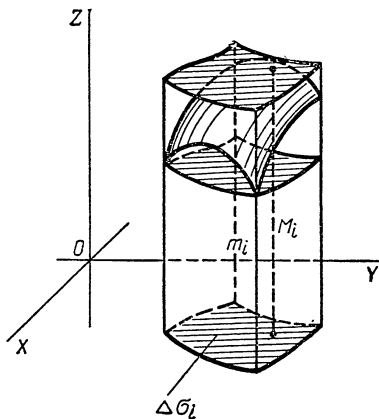


Рис. 366.

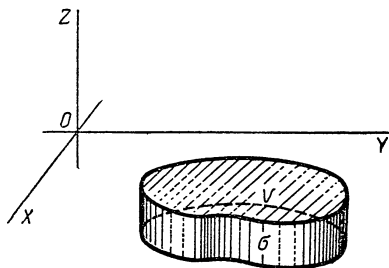


Рис. 367.

Если неположительная функция $f(x, y)$ непрерывна в области σ , то, воспроизводя рассуждения, приведенные в начале главы, получим, что двойной интеграл $\iint_{\sigma} f(x, y) d\sigma$ равен взятому со знаком минус объему цилиндрического тела, ограниченного сверху областью σ плоскости XOY , снизу — поверхностью $z=f(x, y)$, сбоку — цилиндрической поверхностью с образующими, параллельными оси OZ .

Предположим теперь, что функция $f(x, y)$ непрерывна в области σ , причем область σ можно разбить на конечное число частей $\sigma_1, \sigma_2, \dots, \sigma_m$, в каждой из которых функция $f(x, y)$ либо неотрицательна, либо неположительна. Тогда, заменяя двойной интеграл от функции $f(x, y)$ суммой интегралов по областям $\sigma_1, \sigma_2, \dots, \sigma_m$, заключаем, что двойной интеграл $\iint_{\sigma} f(x, y) d\sigma$ равен алгебраической

сумме объемов цилиндрических тел с образующими, параллельными оси OZ , ограниченных с одной стороны частями поверхности $z=f(x, y)$, с другой стороны — соответствующими областями $\sigma_1, \sigma_2, \dots, \sigma_m$ плоскости XOY . В эту сумму объемы тел, лежащих над плоскостью XOY , входят со знаком плюс, а объемы тел, лежащих под плоскостью XOY , входят со знаком минус (рис. 368).

Указанные соображения используются при вычислении объемов тел с помощью двойных интегралов.

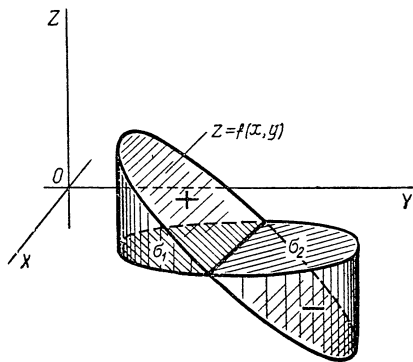


Рис. 368.

§ 9. Теорема о среднем

Так же, как и для определенного интеграла, для двойного интеграла имеет место так называемая **теорема о среднем**.

Теорема. Если функция $f(x, y)$ непрерывна в замкнутой области σ , то в этой области существует точка $P(\xi, \eta)$, такая, что

$$\iint_{\sigma} f(x, y) d\sigma = f(\xi, \eta) \cdot \sigma.$$

Доказательство. Обозначим через m и M соответственно наименьшее и наибольшее значение функции $f(x, y)$ в замкнутой области σ .

Тогда для любой точки (x, y) этой области значения функции $f(x, y)$ удовлетворяют неравенствам

$$m \leq f(x, y) \leq M.$$

Интегрируя эти неравенства (см. свойство 4), будем иметь:

$$\iint_{\sigma} m d\sigma \leq \iint_{\sigma} f(x, y) d\sigma \leq \iint_{\sigma} M d\sigma,$$

или, так как $\iint_{\sigma} m d\sigma = m \iint_{\sigma} d\sigma = m\sigma$,

$$\iint_{\sigma} M d\sigma = M\sigma,$$

$$m\sigma \leq \iint_{\sigma} f(x, y) d\sigma \leq M\sigma.$$

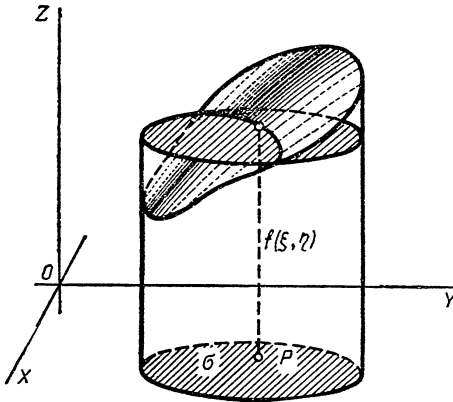


Рис. 369.

Отсюда следует, что число $\frac{\iint_{\sigma} f(x, y) d\sigma}{\sigma}$ заключено между числами m и M :

$$m \leq \frac{\iint_{\sigma} f(x, y) d\sigma}{\sigma} \leq M.$$

Функция $f(x, y)$, непрерывная в замкнутой области σ , между своими наименьшим и наибольшим значениями принимает все промежуточные значения. Следовательно, существует в области σ точка $P(\xi, \eta)$, такая, что

$$\frac{\iint_{\sigma} f(x, y) d\sigma}{\sigma} = f(\xi, \eta),$$

или, что то же,

$$\iint_{\sigma} f(x, y) d\sigma = f(\xi, \eta) \sigma.$$

Для неотрицательной в области σ функции $f(x, y)$ теорема о среднем допускает следующее геометрическое истолкование: объем цилиндрического тела с основанием σ , ограниченного сверху поверхностью $z = f(x, y)$, численно равен объему прямого цилиндра с тем же основанием σ и высотой, равной значению функции $f(x, y)$ в некоторой точке P области σ (рис. 369).

§ 10. Вычисление двойных интегралов

Вычисление двойных интегралов сводится к вычислению определенных интегралов.

Теорема 1. Если функция $f(x, y)$ непрерывна в замкнутой области σ , ограниченной линиями $x = a$, $x = b$ ($a < b$), $y = \varphi(x)$, $y = \psi(x)$ ($\varphi(x)$ и

$\psi(x)$ — непрерывные функции на отрезке $[a, b]$, причем $\varphi(x) \leq \psi(x)$ на этом отрезке) (рис. 370), то имеет место равенство

$$\iint_{\sigma} f(x, y) dx dy = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x, y) dy \right) dx, \quad (1)$$

позволяющее вычисление двойного интеграла свести к последовательному вычислению определенного интеграла от определенного интеграла (или, короче, к вычислению *повторного* интеграла).

Повторный интеграл, стоящий в правой части равенства (1), обычно записывают в виде:

$$\int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy.$$

Доказательство. Рассмотрим n равноотстоящих прямых $x = x_0, x = x_1, \dots, x = x_n$, где $a = x_0 < x_1 < \dots < x_n = b$, и n кривых, заданных уравнениями $y = \varphi_0(x), y = \varphi_1(x), \dots, y = \varphi_n(x)$, где $\varphi_k(x) = \varphi(x) + \frac{k}{n} [\psi(x) - \varphi(x)]$, $k = 0, 1, 2, \dots, n$. Функции $\varphi_k(x)$ непрерывны на отрезке $[a, b]$, причем в каждой точке x этого отрезка

$$\varphi_0(x) = \varphi(x) \leq \varphi_1(x) \leq \varphi_2(x) \leq \dots \leq \varphi_n(x) = \psi(x)$$

(см. рис. 371).

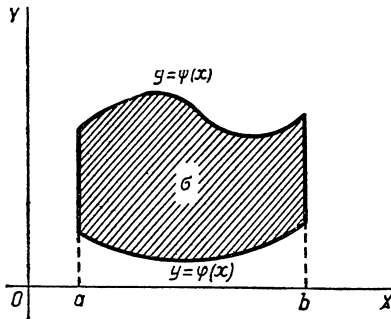


Рис. 370.

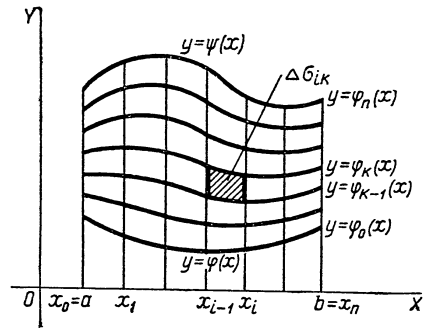


Рис. 371.

Эта система линий разобьет область σ на части. Пусть $\Delta\sigma_{ik}$ — область деления, для которой $x_{i-1} \leq x \leq x_i$, $\varphi_{k-1}(x) \leq y \leq \varphi_k(x)$. Обозначим через m_{ik} и M_{ik} наименьшее и наибольшее значения функции $f(x, y)$ в этой области. Так как $\Delta\sigma_{ik}$ — замкнутая область, эти значения принимаются функцией $f(x, y)$ в некоторых точках области $\Delta\sigma_{ik}$, причем всюду в области $\Delta\sigma_{ik}$ выполняются неравенства $m_{ik} \leq f(x, y) \leq M_{ik}$.

Интегрируя их по y для каждого значения x , $x_{i-1} \leq x \leq x_i$, получаем:

$$m_{ik} [\varphi_k(x) - \varphi_{k-1}(x)] \leq \int_{\varphi_{k-1}(x)}^{\varphi_k(x)} f(x, y) dy \leq M_{ik} [\varphi_k(x) - \varphi_{k-1}(x)]. \quad (2)$$

Стоящий посредине интеграл представляет собой непрерывную функцию параметра x на отрезке $[x_{i-1}, x_i]$. Интегрируя неравенства (2) по x в пределах от x_{i-1} до x_i и учитывая, что $\int_{x_{i-1}}^{x_i} [\varphi_k(x) - \varphi_{k-1}(x)] dx = \Delta\sigma_{ik}$, где $\Delta\sigma_{ik}$ — площадь области $\Delta\sigma_{ik}$, получаем:

$$m_{ik} \Delta\sigma_{ik} \leq \int_{x_{i-1}}^{x_i} dx \int_{\varphi_{k-1}(x)}^{\varphi_k(x)} f(x, y) dy \leq M_{ik} \Delta\sigma_{ik}. \quad (3)$$

Неравенства (3) имеют место в каждой области деления $\Delta\sigma_{ik}$, $i = 1, 2, \dots, n$, $k = 1, 2, \dots, n$. Суммируя их по k от $k = 1$ до $k = n$, находим:

$$\sum_{k=1}^n m_{ik} \Delta\sigma_{ik} \leq \int_{x_{i-1}}^{x_i} dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy \leq \sum_{k=1}^n M_{ik} \Delta\sigma_{ik},$$

так как

$$\begin{aligned} \sum_{k=1}^n \int_{x_{i-1}}^{x_i} dx \int_{\varphi_{k-1}(x)}^{\varphi_k(x)} f(x, y) dy &= \int_{x_{i-1}}^{x_i} \left(\sum_{k=1}^n \int_{\varphi_{k-1}(x)}^{\varphi_k(x)} f(x, y) dy \right) dx = \\ &= \int_{x_{i-1}}^{x_i} dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy. \end{aligned}$$

Суммируя полученные неравенства по i от $i = 1$ до $i = n$, имеем:

$$\sum_{i,k=1}^n m_{ik} \Delta\sigma_{ik} \leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy \leq \sum_{i,k=1}^n M_{ik} \Delta\sigma_{ik},$$

или, что то же,

$$\sum_{i,k=1}^n m_{ik} \Delta\sigma_{ik} \leq \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy \leq \sum_{i,k=1}^n M_{ik} \Delta\sigma_{ik}. \quad (4)$$

Суммы $\sum_{i,k=1}^n m_{ik} \Delta\sigma_{ik}$ и $\sum_{i,k=1}^n M_{ik} \Delta\sigma_{ik}$ являются соответственно *нижней* и *верхней* интегральными суммами для функции $f(x, y)$ и данного разбиения области σ на части. Пусть n стремится к бесконечности. Можно показать, что при этом наибольший из диаметров областей деления $\Delta\sigma_{ik}$ области σ (обозначим его λ) стремится к нулю¹⁾. Так как функция $f(x, y)$ по условию интегрируема в области σ , то при стремлении к нулю шага разбиения λ области σ эти суммы имеют общий предел:

$$\lim_{\lambda \rightarrow 0} \sum_{i,k=1}^n m_{ik} \Delta\sigma_{ik} = \lim_{\lambda \rightarrow 0} \sum_{i,k=1}^n M_{ik} \Delta\sigma_{ik} = \iint_{\sigma} f(x, y) dx dy.$$

Из неравенств (4) следует, что этот предел совпадает с повторным интегралом:

$$\int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy,$$

а значит, равенство (1)

$$\iint_{\sigma} f(x, y) dx dy = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy$$

доказано.

При вычислении двойного интеграла с помощью повторного по формуле (1) сначала вычисляется внутренний интеграл $\int_{\varphi(x)}^{\psi(x)} f(x, y) dy$ при постоянном значении переменной x , в пределах изменения y (для области σ) при постоянном значении x , $a \leq x \leq b$, затем полученная функция от x интегрируется по x в максимальных пределах изменения переменной x для области σ .

¹⁾ На доказательстве этого утверждения мы останавливаться не будем. Заметим лишь, что оно основано на использовании равномерной непрерывности функций $\varphi(x)$ и $\psi(x)$ на отрезке $[a, b]$.

Рассмотрим п р и м е р. Вычислить интеграл $\iint_{\sigma} (x-y) dx dy$, если область σ ограничена линиями $y=0$, $y=x^2$, $x=2$ (рис. 372).

Так как область σ и функция $f(x, y)=x-y$ удовлетворяют условиям теоремы 1, получаем:

$$\begin{aligned}\iint_{\sigma} (x-y) dx dy &= \int_0^2 dx \int_0^{x^2} (x-y) dy = \\ &= \int_0^2 \left(xy - \frac{y^2}{2} \right) \bigg|_{y=0}^{y=x^2} dx = \int_0^2 \left(x^3 - \frac{x^4}{2} \right) dx = \frac{4}{5}.\end{aligned}$$

Поменяв в предыдущих рассуждениях ролями переменные x и y , мы получим следующую теорему.

Теорема 2. Если функция $f(x, y)$ непрерывна в замкнутой области σ , ограниченной линиями $y=c$, $y=d$ ($c < d$), $x=\Phi(y)$, $x=\Psi(y)$ ($\Phi(y)$ и

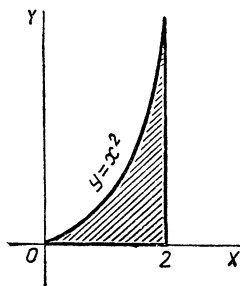


Рис. 372.

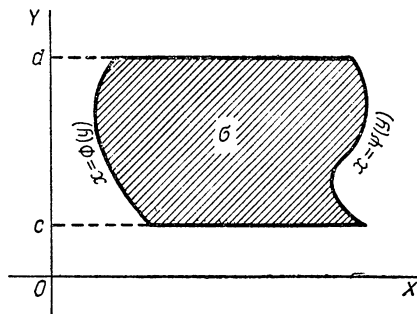


Рис. 373.

$\Psi(y)$ — непрерывные функции на отрезке $[c, d]$, причем всюду на этом отрезке $\Phi(y) \leq \Psi(y)$ (рис. 373), то имеет место равенство

$$\iint_{\sigma} f(x, y) dx dy = \int_c^d dy \int_{\Phi(y)}^{\Psi(y)} f(x, y) dx, \quad (5)$$

позволяющее сводить вычисление двойного интеграла к вычислению повторного интеграла.

При вычислении двойного интеграла с помощью повторного по формуле (5) сначала вычисляется внутренний интеграл

$\int_{\Phi(y)}^{\Psi(y)} f(x, y) dx$ при постоянном значении переменной y , в пределах изменения x (для области σ) при постоянном значении y , $c \leq y \leq d$, затем полученная функция от y интегрируется по y в максимальных пределах изменения y для области σ .

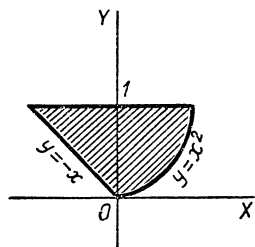


Рис. 374.

П р и м е р. Вычислить интеграл $\iint_{\sigma} x dx dy$ по области σ , ограниченной линиями $y=-x$, $y=1$, $y=x^2$ (см. рис. 374). Так как выполнены условия теоремы 2, применяя формулу (5), получаем:

$$\iint_{\sigma} x dx dy = \int_0^1 dy \int_{-y}^{\sqrt{y}} x dx = \frac{1}{2} \int_0^1 (y - y^3) dy = \frac{1}{2} \left(\frac{y^2}{2} - \frac{y^4}{4} \right) \bigg|_0^1 = \frac{1}{12}.$$

Если функция $f(x, y)$ непрерывна в замкнутой области σ , удовлетворяющей одновременно условиям теорем 1 и 2, то при вычислении двойного интеграла $\iint_{\sigma} f(x, y) d\sigma$ можно выбирать любой порядок интегриро-

вания (внешний интеграл брать по x , внутренний — по y или наоборот). Так, например, если граница области σ каждой прямой, параллельной оси OX , и каждой прямой, параллельной оси OY , пересекается не более чем в двух точках (см. рис. 375), то применимы обе формулы (1) и (5), т. е.

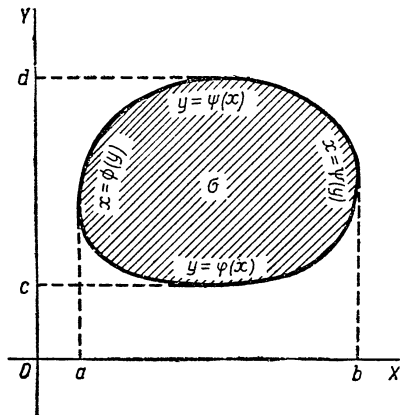


Рис. 375.

$$\begin{aligned}\iint_{\sigma} f(x, y) dx dy &= \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy = \\ &= \int_c^d dy \int_{\Phi(y)}^{\Psi(y)} f(x, y) dx.\end{aligned}$$

Если область σ — прямоугольник со сторонами $x=a$, $x=b$, $y=c$ и $y=d$, а $f(x, y)$ — функция, непрерывная в прямоугольнике $\{a \leq x \leq b, c \leq y \leq d\}$, то, применяя формулы (1) и (5), получаем:

$$\begin{aligned}\iint_{\sigma} f(x, y) dx dy &= \int_a^b dx \int_c^d f(x, y) dy = \\ &= \int_c^d dy \int_a^b f(x, y) dx.\end{aligned}$$

В частности, если $f(x, y) = f_1(x) \cdot f_2(y)$, то двойной интеграл по этому прямоугольнику оказывается равным произведению двух определенных интегралов:

$$\begin{aligned}\iint_{\sigma} f(x, y) dx dy &= \int_a^b dx \int_c^d f_1(x) \cdot f_2(y) dy = \int_a^b f_1(x) dx \int_c^d f_2(y) dy = \\ &= \left(\int_a^b f_1(x) dx \right) \cdot \left(\int_c^d f_2(y) dy \right).\end{aligned}$$

Например, если σ — прямоугольник со сторонами $x = \pm 1$, $y = \pm 2$, то

$$\iint_{\sigma} x^2 e^{1+y} dx dy = \int_{-1}^1 x^2 dx \int_{-2}^2 e^{1+y} dy = \frac{2}{3} (e^3 - e^{-1}).$$

При вычислении двойного интеграла по области σ более сложного вида обычно применяется предварительное разбиение этой области на конечное число частей, удовлетворяющих условиям теорем 1 или 2¹⁾, интеграл по области σ заменяется равной ему суммой интегралов по ее частям, а каждый из полученных интегралов вычисляется по формуле (1) или (5).

Пример. Составить формулу для вычисления интеграла $\iint_{\sigma} f(x, y) dx dy$ по области σ , ограниченной линиями $x^2 + y^2 = 2x$, $x^2 + y^2 = 4x$ (см. рис. 376).

Разобьем область σ на три части σ_1 , σ_2 , σ_3 , как указано на рис. 376. Каждая из полученных областей деления удовлетворяет условиям теоремы 1. При этом область σ_1 ограничена линиями $x=2$, $y = \sqrt{2x-x^2}$, $y = \sqrt{4x-x^2}$, т. е.

$$\iint_{\sigma_1} f(x, y) dx dy = \int_0^2 dx \int_{\sqrt{2x-x^2}}^{\sqrt{4x-x^2}} f(x, y) dy.$$

Область σ_2 ограничена линиями $x=2$, $y = -\sqrt{4x-x^2}$, $y = -\sqrt{2x-x^2}$, т. е.

$$\iint_{\sigma_2} f(x, y) dx dy = \int_0^2 dx \int_{-\sqrt{4x-x^2}}^{-\sqrt{2x-x^2}} f(x, y) dy,$$

¹⁾ Области, которые нельзя разбить на конечное число частей указанного вида, практически встречаются редко и в связи с этим в данном курсе не рассматриваются.

и, наконец, область σ_3 ограничена линиями $x=2, y=-\sqrt{4x-x^2}, y=+\sqrt{4x-x^2}$, т. е.

$$\iint_{\sigma_3} f(x, y) dx dy = \int_2^4 dx \int_{-\sqrt{4x-x^2}}^{\sqrt{4x-x^2}} f(x, y) dy.$$

Складывая эти интегралы, получим:

$$\begin{aligned} \iint_{\sigma} f(x, y) dx dy &= \int_0^2 dx \int_{\sqrt{2x-x^2}}^{\sqrt{4x-x^2}} f(x, y) dy + \int_0^2 dx \int_{-\sqrt{4x-x^2}}^{-\sqrt{2x-x^2}} f(x, y) dy + \\ &+ \int_2^4 dx \int_{-\sqrt{4x-x^2}}^{\sqrt{4x-x^2}} f(x, y) dy. \end{aligned}$$

Приведем в заключение наглядный вывод формулы (1), основанный на геометрической интерпретации двойного интеграла. Этот вывод поможет уяснить геометрическую сторону вопроса перехода от двойного интеграла к повторному. Формула (5) может быть получена аналогичным образом.

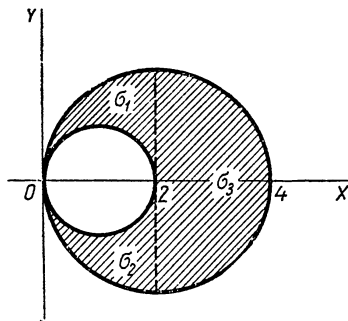


Рис. 376.

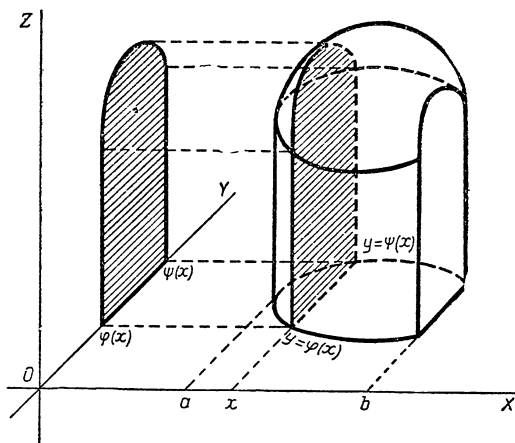


Рис. 377.

Рассмотрим функцию $f(x, y)$, непрерывную и неотрицательную в замкнутой области σ , удовлетворяющей условиям теоремы 1. Тогда задача вычисления двойного интеграла

$$\iint_{\sigma} f(x, y) dx dy$$

равносильна задаче нахождения объема V цилиндрического тела с основанием σ , ограниченного сверху поверхностью $z=f(x, y)$. Если функция $f(x, y)$ такова, что любые два сечения этого тела плоскостями, перпендикулярными оси OX , проектируются на плоскость YOZ одно в другое, то объем тела V может быть найден по формуле (см. гл. VI, разд. 2, § 11)

$$V = \int_a^b S(x) dx. \quad (6)$$

Здесь $S(x)$ — площадь сечения тела плоскостью, перпендикулярной оси OX в точке с абсциссой $x, a \leq x \leq b$. Это сечение представляет собой криволинейную трапецию. Спроектировав ее на плоскость YOZ , получим криволинейную трапецию, ограниченную снизу — отрезком $[\varphi(x), \psi(x)]$ оси OY , сверху — кривой $z=f(x, y)$ (где x — постоянная) (рис. 377). Площадь этой трапеции, равная $S(x)$, находится с помощью определенного интеграла:

$$S(x) = \int_{\varphi(x)}^{\psi(x)} f(x, y) dy.$$

Подставляя значение $S(x)$ в формулу (6), получаем знакомое выражение (1) двойного интеграла через повторный:

$$\iint_{\sigma} f(x, y) d\sigma = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy.$$

Примечание. Это геометрическое доказательство формулы (1) проводится при дополнительных ограничениях на функцию $f(x, y)$. Для произвольной функции $f(x, y)$ доказательство формулы (1) было дано ранее.

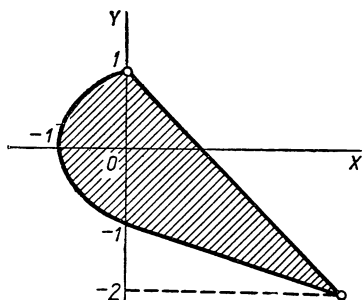


Рис. 378.

§ 11. Некоторые геометрические и механические приложения двойных интегралов

Как уже говорилось выше, двойные интегралы можно применять к вычислению площадей плоских фигур и объемов тел. Рассмотрим несколько примеров.

Пример 1. Найти площадь области σ , ограниченной линиями $y^2 = x + 1$, $x + y = 1$ (рис. 378).

Область σ представляет собой параболический сегмент, ограниченный слева дугой параболы $x = y^2 - 1$, справа — отрезком прямой $x = 1 - y$. Решая совместно уравнения параболы и прямой, находим ординаты точек их пересечения: $y = -2$, $y = +1$.

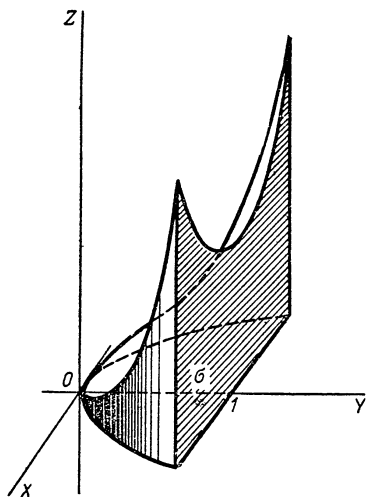


Рис. 379.

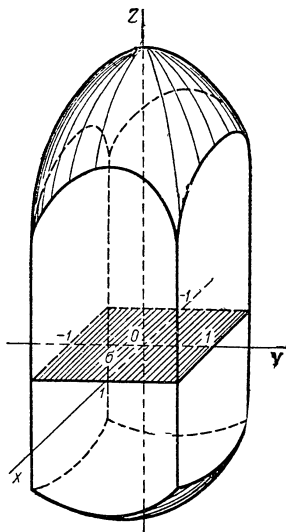


Рис. 380.

Следовательно,

$$\sigma = \iint_{\sigma} dx dy = \int_{-2}^1 dy \int_{y^2-1}^{1-y} dx = \int_{-2}^1 (2 - y - y^2) dy = \frac{9}{2}.$$

Примечание. Если бы мы выбрали обратный порядок интегрирования, область σ предварительно пришлось бы разбить на две части, так как эта область сверху ограничена линией, заданной двумя различными уравнениями (дуга параболы и отрезок прямой).

Пример 2. Найти объем тела V , ограниченного поверхностями $y = x^2$, $y = 1$, $z = 0$, $z = x^2 + y^2$ (рис. 379).

Так как данное тело представляет собой цилиндрическое тело с основанием σ (см. рис. 379), ограниченное сверху параболоидом $z = x^2 + y^2$, то имеем:

$$V = \iint_{\sigma} (x^2 + y^2) dx dy = 2 \int_0^1 dy \int_0^{\sqrt{y}} (x^2 + y^2) dx = 2 \int_0^1 \left(\frac{1}{3} y^{\frac{3}{2}} + y^{\frac{5}{2}} \right) dy = \frac{88}{105}.$$

Пример 3. Найти объем тела V , вырезаемого из бесконечной призмы с гранями $x = \pm 1$, $y = \pm 1$ параболоидами $x^2 + y^2 = 4 - z$, $x^2 + y^2 = 4(z + 2)$ (см. рис. 380).

Объем тела V находим как сумму объемов V_1 и V_2 его частей, лежащих соответственно над и под плоскостью XOY . При этом

$$\begin{aligned} V_1 &= \iint_{\sigma} (4 - x^2 - y^2) dx dy = 4 \int_0^1 dx \int_0^1 (4 - x^2 - y^2) dy = \\ &= 4 \int_0^1 \left(4 - x^2 - \frac{1}{3} \right) dx = \frac{40}{3}, \end{aligned}$$

$$V_2 = - \iint_{\sigma} \left(\frac{x^2 + y^2}{4} - 2 \right) dx dy = -4 \int_0^1 dx \int_0^1 \left(\frac{x^2 + y^2}{4} - 2 \right) dy = \frac{22}{3}.$$

$$V = V_1 + V_2 = \frac{62}{3}.$$

Понятие двойного интеграла можно использовать для определения площадей не только плоских фигур, но и кривых поверхностей.

Будем называть поверхность S , в каждой точке которой определена касательная плоскость и, следовательно, нормаль, *гладкой*, если положение касательной плоскости непрерывно меняется с непрерывным перемещением по поверхности точки касания. Последнее означает, что для любой точки M_0 поверхности и для любого числа $\varepsilon > 0$ существует такая δ -окрестность точки M_0 , что для всех точек M поверхности, лежащих в этой окрестности, углы между нормалью в точках M и M_0 меньше, чем ε .

Пусть, например, поверхность S задана уравнением $z = f(x, y)$, где $f(x, y)$ — функция, имеющая непрерывные частные производные $f'_x(x, y)$ и $f'_y(x, y)$ в замкнутой квадратируемой области σ — проекции поверхности S на плоскость XOY (рис. 381). При таких предположениях

поверхность S имеет непрерывно меняющуюся касательную плоскость и нормаль, т. е. является гладкой.

Для определения площади поверхности поступим следующим образом. Разобьем область σ на n областей $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ ¹⁾ и обозначим через ΔS_i часть поверхности S , проектирующуюся на плоскость XOY в область $\Delta\sigma_i$. При этом поверхность S разобьется на n частей: $\Delta S_1, \Delta S_2, \dots, \Delta S_n$. В каждой области $\Delta\sigma_i$ возьмем произвольно точку $P_i(\xi_i, \eta_i)$, восстановив в этой точке перпендикуляр к плоскости XOY до пересечения с поверхностью S , получим на поверхности ΔS_i точку $M_i(\xi_i, \eta_i, \zeta_i)$. Проведем в точке M_i касательную плоскость к поверхности и рассмотрим ту ее часть ΔT_i , которая на плоскость XOY проектируется в область $\Delta\sigma_i$ (рис. 382). В силу гладкости поверхности вблизи точки M_i касательная плоскость мало отклоняется от поверхности ΔS_i , поэтому естественно считать площадь части ΔS_i поверхности приближенно равной площади части ΔT_i касательной плоскости.

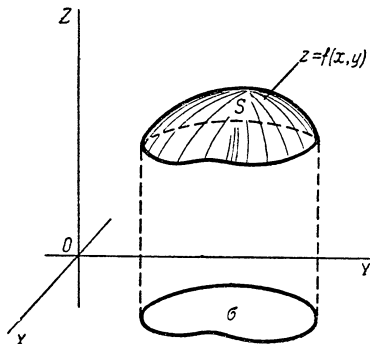


Рис. 381.

¹⁾ Предполагается, что области деления не имеют общих внутренних точек.

Проводя указанные рассуждения для всех областей деления и суммируя результаты, получим приближенное значение площади поверхности S в виде:

$$S \approx \sum_{i=1}^n \Delta T_i.$$

(Здесь ΔT_i обозначает площадь плоской площадки ΔT_i .)

За точное значение *площади поверхности* S по определению принимается число, равное пределу, к которому стремится полученное приближенное значение при стремлении к нулю шага разбиения λ области σ :

$$S = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \Delta T_i.$$

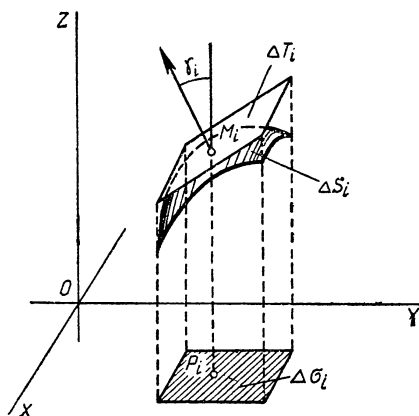


Рис. 382.

Покажем, что в наших предположениях этот предел существует. Обозначим через γ_i острый угол, составляемый с осью OZ нормалью к поверхности S , или, что то же, нормалью к плоскости ΔT_i в точке M_i . Так как область $\Delta \sigma_i$ есть проекция ΔT_i на плоскость XOY , их площади связаны соотношением

$$\Delta T_i \cos \gamma_i = \Delta \sigma_i.$$

Косинус угла γ_i находим, используя уравнение нормали к поверхности $z = f(x, y)$ в точке $M_i(\xi_i, \eta_i, f(\xi_i, \eta_i))$ (см. гл. VII, § 7):

$$\cos \gamma_i = \frac{1}{\sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)}}.$$

Следовательно,

$$\Delta T_i = \frac{\Delta \sigma_i}{\cos \gamma_i} = \sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)} \Delta \sigma_i,$$

а значит,

$$S = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)} \Delta \sigma_i.$$

Под знаком предела стоит интегральная сумма по области σ для функции $\varphi(x, y) = \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)}$. Так как по условию эта функция непрерывна в замкнутой области σ , то указанный предел существует и равен двойному интегралу

$$\iint_{\sigma} \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)} dx dy.$$

Таким образом, мы получили *формулу для вычисления площади поверхности* S :

$$S = \iint_{\sigma} \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)} dx dy. \quad (1)$$

Примечание. Если поверхность S задана уравнением вида $y = f(x, z)$ или $x = f(y, z)$, то за плоскость проекции берется плоскость ZOX

или соответственно YOZ . Соответствующим образом видоизменяются определение площади поверхности и формула (1).

В более сложных случаях при определении площади поверхности предварительно разбивают ее на части рассмотренного выше вида и площадь поверхности S полагают равной сумме площадей ее частей. Поверхность S , имеющую определенную указанным выше образом площадь, называют *квадрируемой*.

Пример. Найти площадь части конической поверхности $z = \sqrt{x^2 + y^2}$, вырезаемой плоскостями $x=0$, $y=0$, $x+y=1$, $x+y=2$ и лежащей в первом октанте (см. рис. 383).

Так как функция $f(x, y) = \sqrt{x^2 + y^2}$ и область σ , являющаяся проекцией данной части поверхности на плоскость XOY , удовлетворяют сформулированным выше условиям, вычисляем площадь поверхности по формуле (1). При этом

$$f'_x(x, y) = \frac{x}{\sqrt{x^2 + y^2}}, \quad f'_y(x, y) = \frac{y}{\sqrt{x^2 + y^2}},$$

$$\sqrt{1 + f'^2_x(x, y) + f'^2_y(x, y)} = \sqrt{2},$$

т. е.

$$S = \iint_{\sigma} \sqrt{2} \, d\sigma = \sqrt{2} \iint_{\sigma} d\sigma = \sqrt{2} \sigma =$$

$$= \frac{3}{2} \sqrt{2} \text{ (1)}.$$

С помощью двойных интегралов можно вычислять ряд механических величин, связанных с непрерывным распределением массы вдоль плоской фигуры. Остановимся подробнее на определении массы, статических моментов, моментов инерции относительно взаимно перпендикулярных осей и положения центра тяжести материальной плоской фигуры (области) σ по известной плотности распределения массы²⁾.

Все эти величины определяются одним и тем же способом: данную область разбивают на конечное число «мелких» частей; делая для каждой области деления упрощающие задачу предположения, находят приближенно искомую величину, а затем, переходя к пределу при неограниченном измельчении областей деления и, следовательно, при неограниченном увеличении числа областей деления, получают точное значение искомой величины.

Пусть, например, известна плотность распределения массы $\mu = \mu(x, y)$ плоской материальной области σ , расположенной в плоскости XOY . Будем считать $\mu(x, y)$ непрерывной функцией в замкнутой области σ . Разобьем область σ на n частей $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$, не имеющих между собой общих внутренних точек. В каждой области деления $\Delta\sigma_i$ возьмем произвольно точку $P_i(\xi_i, \eta_i)$ и предположим, что в пределах области $\Delta\sigma_i$ плотность распределения массы постоянна и равна ее значению в точке P_i . Если диаметр области $\Delta\sigma_i$ мал, то плотность распределения массы в точках этой области, в силу непрерывности функции $\mu(x, y)$, мало отличается от ее значения в точке P_i и, следовательно, масса m_i области $\Delta\sigma_i$ мало отличается от массы $\mu(\xi_i, \eta_i) \Delta\sigma_i$ фиктивной однородной области $\Delta\sigma_i$ с постоянной плотностью $\mu(\xi_i, \eta_i)$. В связи с этим естест-

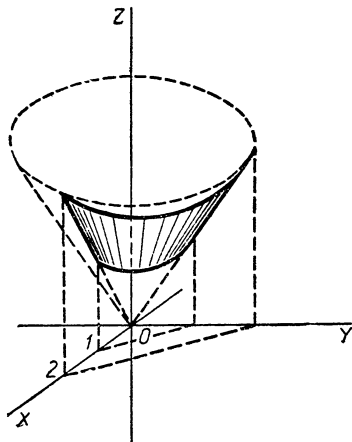


Рис. 383.

¹⁾ Площадь области σ находим как разность площадей прямоугольных треугольников.

²⁾ Средней плотностью распределения массы плоской фигуры называют отношение массы этой фигуры к ее площади. Плотностью распределения массы в точке M данной фигуры называют предел средней плотности распределения массы в части фигуры, содержащей точку M , при стягивании этой части фигуры к точке M .

венно считать: $m_i \approx \mu(\xi_i, \eta_i) \Delta\sigma_i$. Суммируя по всем областям деления, получаем приближенное значение массы m области σ :

$$m \approx \sum_{i=1}^n \mu(\xi_i, \eta_i) \Delta\sigma_i,$$

причем погрешность этого приближенного равенства, вообще говоря, тем меньше, чем меньше диаметры областей деления.

За точное значение массы материальной области σ принимается предел, к которому стремится полученное приближенное значение при стремлении к нулю шага разбиения λ области σ :

$$m = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \mu(\xi_i, \eta_i) \Delta\sigma_i = \iint_{\sigma} \mu(x, y) d\sigma.$$

Примечание. Из полученного равенства следует механический смысл двойного интеграла от непрерывной неотрицательной функции $f(x, y)$. А именно, двойной интеграл $\iint_{\sigma} f(x, y) d\sigma$ численно равен массе плоской

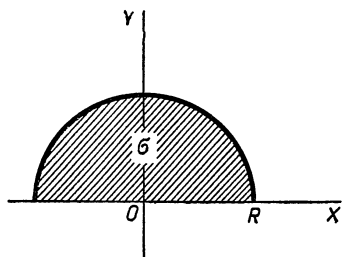


Рис. 384.

материальной области σ , если плотность распределения массы в этой области равна подынтегральной функции: $\mu = f(x, y)$.

Предположим теперь дополнительно, что масса m_i каждой области $\Delta\sigma_i$ сосредоточена в точке $P_i(\xi_i, \eta_i)$, т. е. заменим материальную область σ фиктивной системой из n материальных точек, имеющих (в сумме) ту же массу, что и область σ .

Статические моменты S_x и S_y , моменты инерции J_x , J_y , а также координаты центра тяжести x_c , y_c области σ естественно считать приближенно равными значениям соответствующих величин для указанной системы материальных точек, т. е. считать

$$\begin{aligned} S_x &\approx \sum_{i=1}^n \eta_i m_i \approx \sum_{i=1}^n \eta_i \mu(\xi_i, \eta_i) \Delta\sigma_i, & S_y &\approx \sum_{i=1}^n \xi_i m_i \approx \sum_{i=1}^n \xi_i \mu(\xi_i, \eta_i) \Delta\sigma_i, \\ J_x &\approx \sum_{i=1}^n \eta_i^2 m_i \approx \sum_{i=1}^n \eta_i^2 \mu(\xi_i, \eta_i) \Delta\sigma_i, & J_y &\approx \sum_{i=1}^n \xi_i^2 m_i \approx \sum_{i=1}^n \xi_i^2 \mu(\xi_i, \eta_i) \Delta\sigma_i, \\ x_c &\approx \frac{\sum_{i=1}^n \xi_i m_i}{\sum_{i=1}^n m_i} \approx \frac{\sum_{i=1}^n \xi_i \mu(\xi_i, \eta_i) \Delta\sigma_i}{\sum_{i=1}^n \mu(\xi_i, \eta_i) \Delta\sigma_i}, & y_c &\approx \frac{\sum_{i=1}^n \eta_i m_i}{\sum_{i=1}^n m_i} \approx \frac{\sum_{i=1}^n \eta_i \mu(\xi_i, \eta_i) \Delta\sigma_i}{\sum_{i=1}^n \mu(\xi_i, \eta_i) \Delta\sigma_i}. \end{aligned}$$

Точное значение каждой из этих величин определяем как предел соответствующего приближенного значения при стремлении к нулю шага разбиения λ области σ :

$$\begin{aligned} S_x &= \iint_{\sigma} y \mu(x, y) d\sigma, & S_y &= \iint_{\sigma} x \mu(x, y) d\sigma, \\ J_x &= \iint_{\sigma} y^2 \mu(x, y) d\sigma, & J_y &= \iint_{\sigma} x^2 \mu(x, y) d\sigma, \\ x_c &= \frac{\iint_{\sigma} x \mu(x, y) d\sigma}{\iint_{\sigma} \mu(x, y) d\sigma} = \frac{S_y}{m}, & y_c &= \frac{\iint_{\sigma} y \mu(x, y) d\sigma}{\iint_{\sigma} \mu(x, y) d\sigma} = \frac{S_x}{m}. \end{aligned}$$

В частности, если тело однородное, $\mu = \text{const}$, то

$$x_c = \frac{\iint_{\sigma} x d\sigma}{\sigma}, \quad y_c = \frac{\iint_{\sigma} y d\sigma}{\sigma}.$$

Аналогичным образом получается формула для вычисления полярного момента инерции области σ :

$$J_0 = \iint_{\sigma} (x^2 + y^2) \mu(x, y) d\sigma = J_x + J_y.$$

Вычислим, например, координаты центра тяжести полукруга σ , плотность распределения массы которого пропорциональна квадрату расстояния точки от диаметра этого полукруга.

Пусть ось OX проходит через диаметр полукруга σ , а ось OY перпендикулярна оси OX в его центре O (см. рис. 384). При таком выборе осей координат плотность $\mu(x, y) = ky^2$ (k — коэффициент пропорциональности), $x_c = 0$, $y_c = \frac{S_x}{m}$, где

$$S_x = \iint_{\sigma} ky^3 dx dy = k \int_{-R}^R dx \int_0^{\sqrt{R^2 - x^2}} y^3 dy = \frac{4}{15} kR^5,$$

$$m = \iint_{\sigma} ky^2 dx dy = k \int_{-R}^R dx \int_0^{\sqrt{R^2 - x^2}} y^2 dy = \frac{\pi}{8} kR^4.$$

Следовательно,

$$x_c = 0, \quad y_c = \frac{32}{15\pi} R.$$

§ 12. Формула Грина — Остроградского

Формула Грина — Остроградского устанавливает связь между двойными и криволинейными интегралами (второго рода).

Рассмотрим в плоскости XOY область σ , ограниченную контуром $L(ABCD)$, состоящим из кривых $y = \varphi(x)$, $y = \psi(x)$ [$\varphi(x)$ и $\psi(x)$ — непрерывные функции x , $a \leq x \leq b$, причем $\varphi(x) \leq \psi(x)$] и отрезков прямых $x = a$, $x = b$ (см. рис. 385). В частности, отрезки прямых (AD и BC) могут вырождаться в точки.

Пусть функции $P(x, y)$ и ее частная производная $\frac{\partial P(x, y)}{\partial y}$ непрерывны в замкнутой области σ . При этих условиях существует двойной интеграл $\iint_{\sigma} \frac{\partial P(x, y)}{\partial y} dx dy$, причем

$$\iint_{\sigma} \frac{\partial P(x, y)}{\partial y} dx dy = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} \frac{\partial P(x, y)}{\partial y} dy.$$

Внутренний интеграл вычисляется при постоянном значении x , $a \leq x \leq b$, по формуле Ньютона — Лейбница:

$$\int_{\varphi(x)}^{\psi(x)} \frac{\partial P(x, y)}{\partial y} dy = P(x, y) \Big|_{y=\varphi(x)}^{y=\psi(x)} = P(x, \psi(x)) - P(x, \varphi(x)).$$

Следовательно,

$$\iint_{\sigma} \frac{\partial P(x, y)}{\partial y} dx dy = \int_a^b P(x, \psi(x)) dx - \int_a^b P(x, \varphi(x)) dx.$$

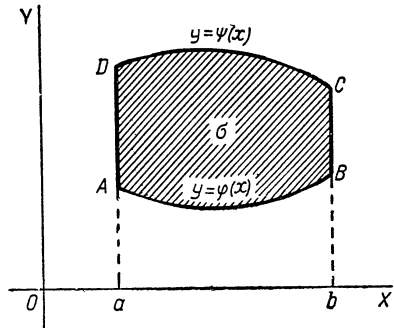


Рис. 385.

Интеграл $\int_a^b P(x, \psi(x)) dx$ может быть заменен равным ему криволинейным интегралом $\int_{DC} P(x, y) dx$ (см. § 3):

$$\int_a^b P(x, \psi(x)) dx = \int_{DC} P(x, y) dx.$$

Аналогично

$$\int_a^b P(x, \varphi(x)) dx = \int_{AP} P(x, y) dx.$$

Таким образом,

$$\iint_{\sigma} \frac{\partial P(x, y)}{\partial y} dx dy = \int_{BC} P(x, y) dx - \int_{AB} P(x, y) dx.$$

Вычитая два равных нулю интеграла $\int_{BC} P(x, y) dx$ и $\int_{DA} P(x, y) dx$ по отрезкам BC и DA , перпендикулярным оси OX , получим:

$$\begin{aligned} \iint_{\sigma} \frac{\partial P(x, y)}{\partial y} dx dy &= \int_{BC} P(x, y) dx - \int_{AB} P(x, y) dx - \int_{BC} P(x, y) dx - \\ &- \int_{DA} P(x, y) dx = - \left[\int_{AB} P(x, y) dx + \int_{BC} P(x, y) dx + \int_{CD} P(x, y) dx + \right. \\ &\quad \left. + \int_{DA} P(x, y) dx \right]. \end{aligned}$$

В квадратных скобках имеем криволинейный интеграл по контуру L области σ в положительном направлении¹⁾, т. е.

$$\iint_{\sigma} \frac{\partial P(x, y)}{\partial y} dx dy = - \oint_L P(x, y) dx. \quad (1)$$

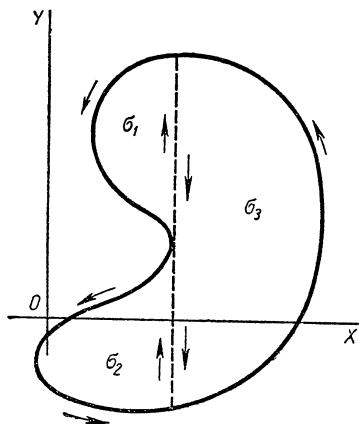


Рис. 386.

Формула (1), выражающая двойной интеграл по области σ через криволинейный интеграл по контуру L этой области, выведена для областей специального вида. Легко видеть, что эта формула остается в силе для любой области σ , которую можно прямыми, параллельными оси OY , разбить на конечное число частей указанного выше вида (рис. 386). Действительно, составляя для каждой области деления равенство (1) и складывая почленно полученные равенства, будем иметь слева — двойной интеграл по всей об-

ласти σ , а справа — взятый со знаком минус криволинейный интеграл по контуру L области σ , так как интегралы по всем вспомогательным отрезкам берутся дважды во взаимно-противоположных направлениях и в сумме равны нулю.

Аналогично устанавливается формула

$$\iint_{\sigma} \frac{\partial Q(x, y)}{\partial x} dx dy = \oint_L Q(x, y) dy \quad (2)$$

¹⁾ Напомним, что под положительным направлением обхода контура лежащей в плоскости XOY области σ понимается направление обхода контура, при котором ограниченная им область σ остается слева.

для случая, когда функция $Q(x, y)$ и ее частная производная $\frac{\partial Q}{\partial x}$ непрерывны в замкнутой области σ , ограниченной кривыми $x = \Phi(y)$, $x = \Psi(y)$ ($\Phi(y)$ и $\Psi(y)$ — непрерывные функции y для $c \leq y \leq d$, $\Phi(y) \leq \Psi(y)$) и отрезками прямых $y = c$ и $y = d$ (рис. 387).

Формула (2) также справедлива для любой области σ , которую можно прямыми, параллельными оси OX , разбить на конечное число частей указанного вида.

Пусть теперь область σ одновременно удовлетворяет условиям, при которых справедливы формулы (1) и (2), а функции $P(x, y)$, $\frac{\partial P}{\partial y}$, $Q(x, y)$ и $\frac{\partial Q}{\partial x}$ непрерывны в замкнутой области σ . Тогда имеют место формулы (1) и (2). Вычитая почленно (2) из (1), получим формулу

$$\iint_{\sigma} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \oint_L P dx + Q dy. \quad (3)$$

Эта важная формула, устанавливающая связь между двойным интегралом по области σ и криволинейным интегралом по ее контуру L , называется *формулой Грина—Остроградского*.

Формулы (1) и (2) можно рассматривать как ее частные случаи.

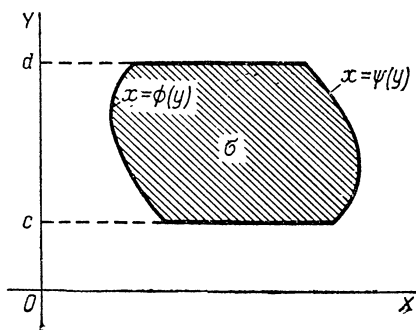


Рис. 387.

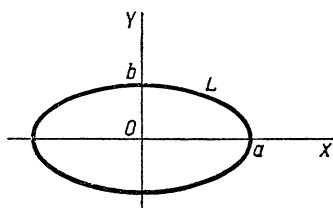


Рис. 388.

Примечание. Можно доказать, что формула Грина—Остроградского имеет место для любой области σ , ограниченной кусочно-гладким контуром L .

Применим формулу Грина—Остроградского к вычислению площади области σ указанного выше вида. Рассмотрим функции $P(x, y) = -y$, $Q(x, y) = 0$. Тогда $\frac{\partial P}{\partial y} = -1$, $\frac{\partial Q}{\partial x} = 0$. По формуле (3) имеем:

$$\iint_{\sigma} (0 + 1) dx dy = \oint_L -y dx + 0 dy.$$

Интеграл $\iint_{\sigma} dx dy$ равен площади области σ , следовательно,

$$\sigma = - \oint_L y dx.$$

Полагая $P(x, y) = 0$, $Q(x, y) = x$, аналогичным образом находим, что

$$\sigma = \oint_L x dy.$$

И, наконец, взяв функции $P(x, y) = -\frac{1}{2}y$, $Q(x, y) = \frac{1}{2}x$, получаем формулу

$$\sigma = \frac{1}{2} \oint_L x dy - y dx.$$

Таким образом, вычисление площади плоской области σ с помощью криволинейного интеграла по ее контуру можно производить по любой из следующих формул:

$$\sigma = \oint_L -y dx = \oint_L x dy = \frac{1}{2} \oint_L x dy - y dx.$$

Вычислим, например, площадь, ограниченную эллипсом $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, по формуле $\sigma = \oint_L x dy$ (рис. 388). Используя параметрические уравнения эллипса: $x = a \cos t$, $y = b \sin t$, $0 \leq t \leq 2\pi$, получаем:

$$\sigma = \oint_L x dy = \int_0^{2\pi} a \cos t \cdot b \cos t dt = \frac{ab}{2} \int_0^{2\pi} (1 + \cos 2t) dt = \pi ab.$$

§ 13. Замена переменных в двойном интеграле

Якобиан и его геометрический смысл. Пусть даны две плоскости с выбранными на них прямоугольными декартовыми системами координат UOV и XOY . Рассмотрим две (ограниченные или неограниченные) области ω и σ , лежащие соответственно в плоскостях UOV и XOY , и предположим, что функции

$$\begin{aligned} x &= \varphi(u, v), \\ y &= \psi(u, v) \end{aligned} \quad (1)$$

устанавливают *взаимно однозначное соответствие* между точками этих областей. Это значит, что каждой точке (u_0, v_0) области ω соответствует одна определенная точка (x_0, y_0) , где $x_0 = \varphi(u_0, v_0)$, $y_0 = \psi(u_0, v_0)$ области σ , называемая *образом* точки (u_0, v_0) , и, наоборот, каждой точке (x_0, y_0) области σ соответствует одна определенная точка (u_0, v_0) области ω [именно та точка, образом которой является точка (x_0, y_0)]. Будем говорить, что формулы (1) задают *взаимно однозначное отображение* области ω на область σ (и наоборот).

Пусть функции $\varphi(u, v)$ и $\psi(u, v)$ непрерывны в области ω вместе со своими частными производными первого порядка. Тогда определитель

$$\begin{vmatrix} \varphi'_u(u, v) & \varphi'_v(u, v) \\ \psi'_u(u, v) & \psi'_v(u, v) \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

будет непрерывной функцией переменных u и v , определенной в области ω . Этот функциональный определитель, называемый *определителем Якоби*, или *якобианом отображения* (1), принято обозначать $J(u, v)$ или символом $\frac{\partial(x, y)}{\partial(u, v)}$, похожим на символ, обозначающий производную¹⁾.

Предположим, что в области ω якобиан отображения (1) отличен от нуля (и, следовательно, в силу непрерывности сохраняет свой знак). Можно показать, что при отображении (1) линия будет отображаться в линию, область в область. В частности, если точка (u, v) , в области ω опишет замкнутый кусочно-гладкий контур в положительном направлении, то ее образ — точка (x, y) в области σ — также опишет замкнутый кусочно-гладкий контур в положительном или отрицательном направлении.

Рассмотрим в области ω любую область ω' , ограниченную кусочно-гладким контуром Γ' . Пусть $u = u(t)$, $v = v(t)$, где $\alpha \leq t \leq \beta$ — параме-

¹⁾ Смысл этого обозначения станет яснее в конце параграфа.

трические уравнения контура Γ^1), причем положительному направлению обхода контура Γ' соответствует изменение параметра t от α к β . Обозначим через σ' область, в которую отображается на плоскость XOY область ω' , а через L' — кусочно-гладкий контур области σ' (линия L' получается в результате отображения линии Γ') (рис. 389). Параметрические уравнения линии L' имеют вид: $x = \varphi(u(t), v(t))$, $y = \psi(u(t), v(t))$, т. е. $x = x(t)$, $y = y(t)$, $\alpha \leq t \leq \beta$, причем изменению параметра t от α к β может соответствовать как положительное, так и отрицательное направление обхода контура L' .

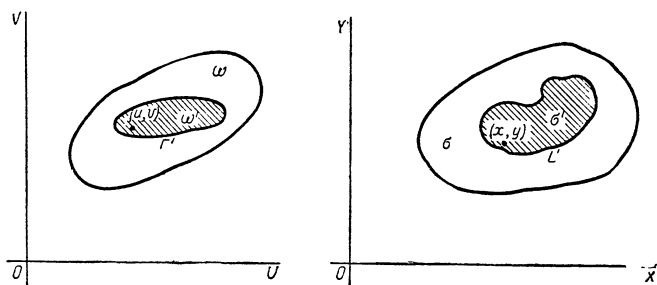


Рис. 389.

Выясним, как при отображении (1) изменяется площадь области ω' . Будем исходить из формулы, выражающей площадь области σ' через криволинейный интеграл по контуру L' этой области:

$$\sigma' = \oint_{L'} x dy.$$

Применяя правило для вычисления криволинейного интеграла, находим, что

$$\sigma' = \oint_{L'} x dy = \pm \int_{\alpha}^{\beta} x(t) y'(t) dt$$

или, так как $y'(t) = \frac{\partial y}{\partial u} u'(t) + \frac{\partial y}{\partial v} v'(t)$, то

$$\sigma' = \pm \int_{\alpha}^{\beta} [x(t) \frac{\partial y}{\partial u} u'(t) + x(t) \frac{\partial y}{\partial v} v'(t)] dt.$$

Знак плюс имеет место, если положительному направлению обхода контура Γ' при отображении (1) соответствует положительное направление обхода контура L' и знак минус в противном случае.

Полученный интеграл равен криволинейному интегралу:

$$\oint_{\Gamma'} x \frac{\partial y}{\partial u} du + x \frac{\partial y}{\partial v} dv.$$

Следовательно,

$$\sigma' = \pm \oint_{\Gamma'} x \frac{\partial y}{\partial u} du + x \frac{\partial y}{\partial v} dv.$$

¹⁾ Согласно замечанию, приведенному в сноске ¹⁾ на стр. 509, функции $u(t)$ и $v(t)$ можно выбрать так, что их производные $u'(t)$ и $v'(t)$ будут на отрезке $[\alpha, \beta]$ непрерывны всюду, кроме, может быть, конечного числа точек разрыва 1 рода.

Интеграл по контуру Γ по формуле Грина — Остроградского преобразуем в двойной интеграл по области ω' :

$$\begin{aligned}\sigma' &= \pm \iint_{\omega'} \left[\frac{\partial}{\partial u} \left(x \frac{\partial y}{\partial v} \right) - \frac{\partial}{\partial v} \left(x \frac{\partial y}{\partial u} \right) \right] du dv = \\ &= \pm \iint_{\omega'} \left[\frac{\partial x}{\partial u} \frac{\partial y}{\partial v} + x \frac{\partial^2 y}{\partial v \partial u} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} - x \frac{\partial^2 y}{\partial u \partial v} \right] du dv.\end{aligned}$$

Предполагая, что у функции $y = \psi(u, v)$ смешанные частные производные второго порядка в области ω непрерывны, а значит, не зависят от последовательности дифференцирований, находим:

$$\sigma' = \pm \iint_{\omega'} \left[\frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \right] du dv,$$

или, что то же,

$$\sigma' = \pm \iint_{\omega'} J(u, v) du dv. \quad (2)$$

Применяя к формуле (2) теорему о среднем, получаем соотношение

$$\sigma' = \pm J(\mu, \nu) \omega', \quad (3)$$

где ω' — площадь области ω' , а (μ, ν) — некоторая точка этой области.

Так как площади σ' и ω' больше нуля, равенство (3) можно переписать в виде:

$$\sigma' = |J(\mu, \nu)| \omega'. \quad (4)$$

Соотношение (4) показывает, что коэффициент изменения площади области при отображении (1) равен абсолютной величине якобиана отображения в некоторой точке этой области:

$$\frac{\sigma'}{\omega'} = |J(\mu, \nu)|, \quad (5)$$

где (μ, ν) — точка области ω' .

На основании полученных результатов можно сделать некоторые заключения о величине и знаке якобиана отображения (1).

Пусть (u, v) — любая точка области ω , ω' — область (с кусочно-гладкой границей), содержащая внутри точку (u, v) , а σ' — ее образ при отображении (1). Переходя к пределу в соотношении (5) при стремлении к нулю диаметра области ω' и, следовательно, при стягивании области ω' к точке (u, v) , в силу непрерывности якобиана $J(u, v)$ получим равенство

$$\lim \frac{\sigma'}{\omega'} = |J(u, v)|,$$

из которого следует, что модуль якобиана отображения (1) в любой точке области ω равен коэффициенту изменения площади бесконечно малой области, содержащей эту точку, при данном отображении.

Вернемся теперь к соотношению (2):

$$\sigma' = \pm \iint_{\omega'} J(u, v) du dv.$$

По предположению якобиан $J(u, v)$ в области ω сохраняет свой знак, следовательно, знак интеграла $\iint_{\omega'} J(u, v) du dv$ совпадает со знаком яко-

биана $J(u, v)$. Так как величина площади σ' больше нуля, то знак плюс перед двойным интегралом берется, если $J(u, v)$ в области ω больше нуля и знак минус — в противном случае. Учитывая сделанные ранее замечания о знаке перед интегралом в формуле (2), заключаем, что если в области ω якобиан отображения (1) положителен, то при отображе-

нии (1) положительному направлению обхода любого контура соответствует положительное направление обхода его образа, если же в области ω якобиан отображения (1) отрицателен, то положительному направлению обхода контура соответствует отрицательное направление обхода его образа.

Примечание. Якобиан отображения (1) можно рассматривать как некоторый двумерный аналог обыкновенной производной функции одной переменной.

Пусть функция $y=f(x)$ непрерывна и дифференцируема на интервале (a, b) , причем производная $f'(x)$ сохраняет на этом интервале свой знак. Эта функция устанавливает взаимно однозначное соответствие между точками интервала (a, b) оси OY и точками некоторого интервала (конечного или бесконечного) оси OY . Рассмотрим любой отрезок $[\alpha, \beta]$ на интервале (a, b) . По теореме Лагранжа:

$$f(\beta) - f(\alpha) = f'(\xi)(\beta - \alpha),$$

где ξ — точка отрезка $[\alpha, \beta]$.

Обозначив через δ длину отрезка $[\alpha, \beta]$, а через d — длину отрезка, соответствующего ему на оси OY ¹⁾, запишем теорему Лагранжа в виде: $d = |f'(\xi)| \cdot \delta$, откуда следует, что

$$\frac{d}{\delta} = |f'(\xi)|,$$

т. е. коэффициент изменения длины отрезка при отображении $y=f(x)$ равен модулю производной в некоторой точке отрезка.

Стягивая отрезок $[\alpha, \beta]$ к некоторой фиксированной точке x отрезка, в пределе [если $f'(x)$ непрерывна] получим соотношение $|f'(x)| = \lim_{\delta \rightarrow 0} \frac{d}{\delta}$.

Соотношение (6) является его обобщением на двухмерный случай.

Вывод формулы замены переменных в двойном интеграле. Рассмотрим двойной интеграл

$$\iint_{\sigma} f(x, y) dx dy \quad (6)$$

от функции $f(x, y)$, непрерывной в замкнутой области σ , ограниченной кусочно-гладким контуром.

Предположим, что формулы

$$\begin{aligned} x &= \varphi(u, v), \\ y &= \psi(u, v) \end{aligned} \quad (7)$$

задают взаимно однозначное отображение замкнутой области ω плоскости UOV на замкнутую область σ плоскости XOY , причем это отображение удовлетворяет всем условиям, указанным в начале параграфа.

Поставим своей целью заменить двойной интеграл (6) по переменным x и y (по области σ) равным ему двойным интегралом по переменным u и v (по области ω). Для этого разобьем область ω с помощью кусочно-гладких кривых²⁾ на n областей $\Delta\omega_1, \Delta\omega_2, \dots, \Delta\omega_n$. Образы этих кривых на плоскости XOY , в свою очередь, будут кусочно-гладкими кривыми. Они разобьют область σ на n областей $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$. Площади соответствующих областей $\Delta\omega_i$ и $\Delta\sigma_i$ связаны соотношением

$$\Delta\sigma_i = |J(\mu_i, \nu_i)| \Delta\omega_i,$$

где $M_i(\mu_i, \nu_i)$ — некоторая точка области $\Delta\omega_i$, а $J(\mu_i, \nu_i)$ — якобиан отображения (7) в этой точке (рис. 390). Обозначим через $P_i(\xi_i, \eta_i)$ лежащий в области $\Delta\sigma_i$ образ точки $M_i(\mu_i, \nu_i)$ при отображении (7):

$$\xi_i = \varphi(\mu_i, \nu_i), \quad \eta_i = \psi(\mu_i, \nu_i), \quad i = 1, 2, \dots, n.$$

¹⁾ $d = |f(\beta) - f(\alpha)|$.

²⁾ Например, с помощью прямых, параллельных координатным осям.

Для полученного разбиения области σ на части $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ и специального выбора точек $P_1, P_2, \dots, P_n$, на этих частях составляем интегральную сумму для двойного интеграла (6):

$$\sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i.$$

По построению имеем:

$$\sum_{i=1}^n f(\xi_i, \eta_i) \Delta\sigma_i = \sum_{i=1}^n f[\varphi(\mu_i, \nu_i), \psi(\mu_i, \nu_i)] |J(\mu_i, \nu_i)| \Delta\omega_i. \quad (8)$$

В правой части этого равенства стоит интегральная сумма для двойного интеграла по области ω от непрерывной функции

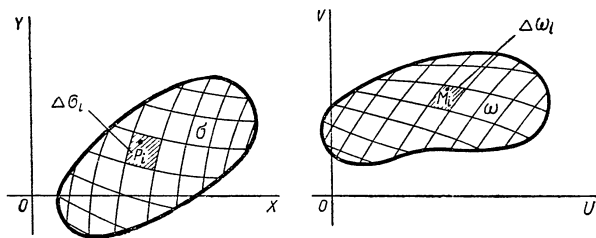


Рис. 390.

$f(\varphi(u, v), \psi(u, v)) |J(u, v)|$. При стремлении к нулю шага разбиения λ_ω области ω стремится к нулю шаг разбиения λ_σ области σ ¹⁾.

Переходя к пределу в равенстве (3), получаем:

$$\iint_{\sigma} f(x, y) dx dy = \iint_{\omega} f(\varphi(u, v), \psi(u, v)) |J(u, v)| du dv. \quad (9)$$

Эта формула и называется *формулой замены переменных в двойном интеграле*.

Примечание. Можно показать, что формула (9) имеет место также и в том случае, если отображение (7) области ω на область σ будет взаимно однозначным лишь для внутренних точек этих областей.

Пример. Найти площадь криволинейного четырехугольника, ограниченного линиями $y=x, y=2x, y=x^2, y=4x^2$ (рис. 391). Формулы $\frac{y}{x}=u, \frac{y}{x^2}=v$ или, что то же, $x=\frac{u}{v}, y=\frac{u^2}{v}$ осуществляют взаимно однозначное отображение данного криволинейного четырехугольника σ плоскости XOY на прямоугольник со сторонами $u=1, u=2, v=1, v=4$ плоскости UOV .

Якобиан отображения

$$J(u, v) = \begin{vmatrix} \frac{1}{v} & -\frac{u}{v^2} \\ \frac{2u}{v} & -\frac{u^2}{v^2} \end{vmatrix} = \frac{u^2}{v^3}.$$

В области ω имеем: $J(u, v) = \frac{u^2}{v^3} > 0$. Следовательно,

$$\sigma = \iint_{\sigma} dx dy = \iint_{\omega} \frac{u^2}{v^3} du dv = \int_1^2 u^2 du \int_1^4 \frac{dv}{v^3} = \frac{35}{32}.$$

¹⁾ Это вытекает из равномерной непрерывности функций $\varphi(u, v)$ и $\psi(u, v)$ в замкнутой области ω .

Переход в двойном интеграле к полярным координатам. Положение точки $M(x, y)$ плоскости XOY можно однозначно определить заданием пары чисел ρ и φ , где ρ — длина радиуса-вектора $\overline{OM}$, а φ — угол, составляемый этим радиусом-вектором с осью OX (φ — полярный угол точки M) (рис. 392). Числа ρ и φ , как уже говорилось ранее, называются *полярными координатами* точки M . Они связаны с ее декартовыми координатами x и y соотношениями:

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi. \quad (10)$$

Будем рассматривать

$$0 \leq \rho < +\infty, \quad 0 \leq \varphi \leq 2\pi.$$

На формулы $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, где переменные ρ и φ изменяются в указанных выше пределах, можно смотреть как на задание отображения «полосы» $T\{0 \leq \rho < +\infty, 0 \leq \varphi \leq 2\pi\}$ (рис. 393) плоскости с прямоугольной системой координат $\rho O\varphi$ на всю плоскость XOY . Это отображение взаимно однозначное всюду, кроме границы полосы T . Действительно, каждой внутренней точке (ρ, φ) , $0 < \rho < +\infty, 0 < \varphi < 2\pi$, полосы T соответствует одна точка (x, y) (где $x = \rho \cos \varphi, y = \rho \sin \varphi$) плоскости XOY . Эта точка отлична от начала координат O (так как $\rho \neq 0$) и не лежит на положительной полуоси OX (так как $\varphi \neq 0$ и $\varphi \neq 2\pi$). Наоборот, каждой точке (x, y) плоскости XOY , отличной от начала O и не лежащей на положительной полуоси OX , соответствует, и притом одна, точка (ρ, φ) полосы T [именно та точка, у которой ρ и φ — полярные координаты точки (x, y)]. Границе полосы T при отображении (10) соответствуют начало координат O и выходящая из него положительная полуось OX . При этом все точки $(0, \varphi)$, $0 \leq \varphi \leq 2\pi$, плоскости $\rho O\varphi$ отображаются в одну точку $O(0, 0)$ плоскости XOY , а каждые две

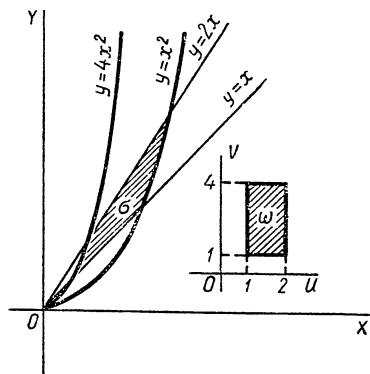


Рис. 391.

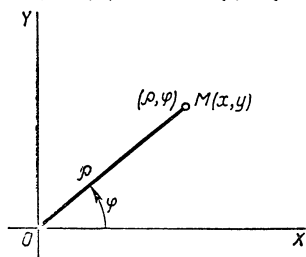


Рис. 392.

точка $(\rho, 0)$ и $(\rho, 2\pi)$, $0 < \rho < +\infty$, плоскости $\rho O\varphi$ отображаются в одну и ту же точку $(x = \rho, y = 0)$ положительной полуоси OX .

Вычислим якобиан отображения (10):

$$J(\rho, \varphi) = \frac{\partial(x, y)}{\partial(\rho, \varphi)} = \begin{vmatrix} \cos \varphi & -\rho \sin \varphi \\ \sin \varphi & \rho \cos \varphi \end{vmatrix} = \rho.$$

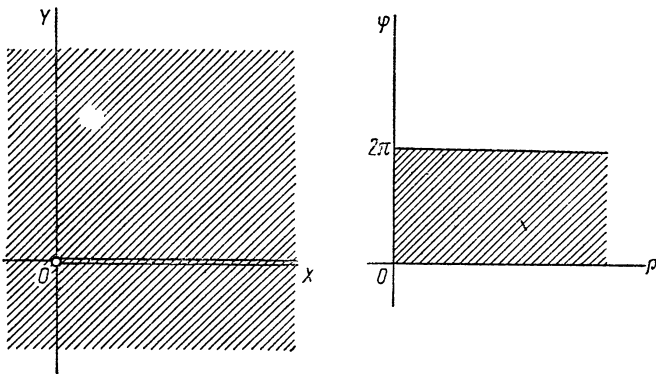


Рис. 393.

точка $(\rho, 0)$ и $(\rho, 2\pi)$, $0 < \rho < +\infty$, плоскости $\rho O\varphi$ отображаются в одну и ту же точку $(x = \rho, y = 0)$ положительной полуоси OX .

Вычислим якобиан отображения (10):

В каждой точке полосы T [кроме граничных точек $(0, \varphi)$, $0 \leq \varphi \leq 2\pi$] имеем:

$$J(\rho, \varphi) = \rho > 0.$$

Применяя выведенную выше общую формулу замены переменных в двойном интеграле к отображению (10) (роль переменных u и v играют полярные координаты ρ и φ), получим так называемую *формулу перехода в двойном интеграле к полярным координатам*:

$$\iint_{\sigma} f(x, y) dx dy = \iint_{\omega} f(\rho \cos \varphi, \rho \sin \varphi) \rho d\rho d\varphi. \quad (11)$$

Здесь σ — данная область плоскости XOY , а ω — область плоскости $\rho O\varphi$ ¹⁾, образом которой при отображении (10) является область σ .

Из сделанного на стр. 544 замечания следует, что нарушение взаимной однозначности отображения (10) в указанных выше точках не влияет на справедливость формулы (11). Переход в двойном интеграле к полярным координатам обычно удобно использовать в тех случаях,

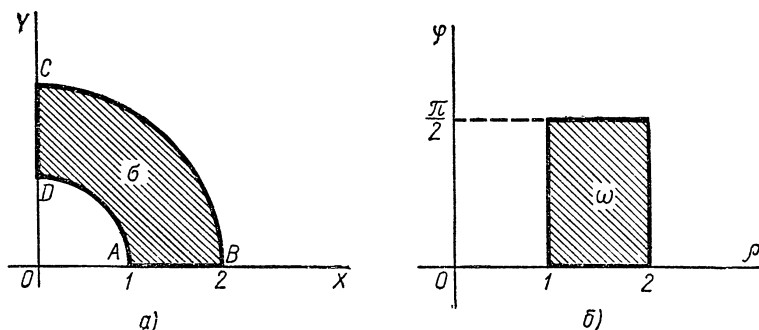


Рис. 394.

когда подынтегральная функция зависит от $x^2 + y^2$ или от $\operatorname{arctg} \frac{y}{x}$, а также в тех случаях, когда граница области σ содержит дуги окружностей и лучи, выходящие из начала координат.

Вычислим, например, интеграл $\iint_{\sigma} e^{x^2 + y^2} dx dy$ по четверти кольца $1 \leq x^2 + y^2 \leq 4$, лежащей в первом квадранте (рис. 394, а). Производя замену переменных по формулам $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, получаем $\iint_{\sigma} e^{x^2 + y^2} dx dy = \iint_{\omega} e^{\rho^2} \rho d\rho d\varphi$, где ω — область плоскости $\rho O\varphi$, образом которой является данная четверть кольца.

Так как при отображении $x = \rho \cos \varphi$, $y = \rho \sin \varphi$ дуге AD окружности $x^2 + y^2 = 1$ соответствует отрезок прямой $\rho = 1$, $0 \leq \varphi \leq \frac{\pi}{2}$, дуге BC окружности $x^2 + y^2 = 4$ — отрезок прямой $\rho = 2$, $0 \leq \varphi \leq \frac{\pi}{2}$, отрезкам AB ($y = 0$) и CD ($x = 0$) — отрезки прямых $\varphi = 0$ и $\varphi = \frac{\pi}{2}$ ($1 \leq \rho \leq 2$), то область ω представляет собой прямоугольник $\{1 \leq \rho \leq 2, 0 \leq \varphi \leq \frac{\pi}{2}\}$ в плоскости $\rho O\varphi$ (рис. 394, б).

Переходя от двойного интеграла по переменным ρ и φ к повторному, получаем:

$$\iint_{\sigma} e^{x^2 + y^2} dx dy = \int_0^{\frac{\pi}{2}} d\varphi \int_1^2 e^{\rho^2} \rho d\rho = \frac{\pi}{4} (e^4 - e).$$

Расстановку пределов при вычислении двойного интеграла в полярных координатах можно производить, используя чертеж области σ на плоскости XOY и геометрический смысл полярных координат.

¹⁾ Точнее, область, лежащая в полосе T .

Пусть, например, область σ ограничена двумя кривыми, полярные уравнения которых $\rho = h(\varphi)$ и $\rho = H(\varphi)$ ($h(\varphi)$ и $H(\varphi)$ — непрерывные функции φ , $\alpha \leq \varphi \leq \beta$, причем $h(\varphi) \leq H(\varphi)$), и двумя полупрямыми, выходящими из начала O под углами $\varphi = \alpha$ и $\varphi = \beta$ к оси OX (см. рис. 395, а).

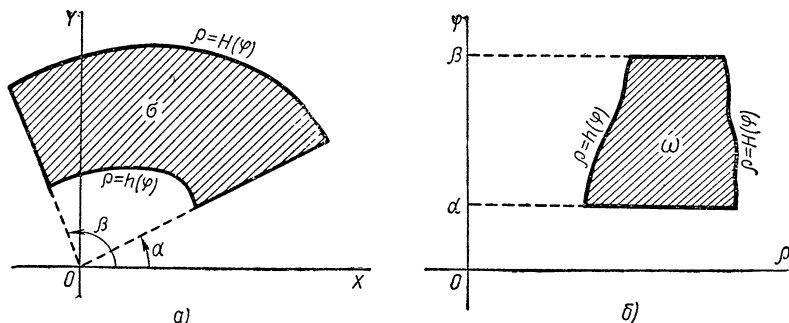


Рис. 395.

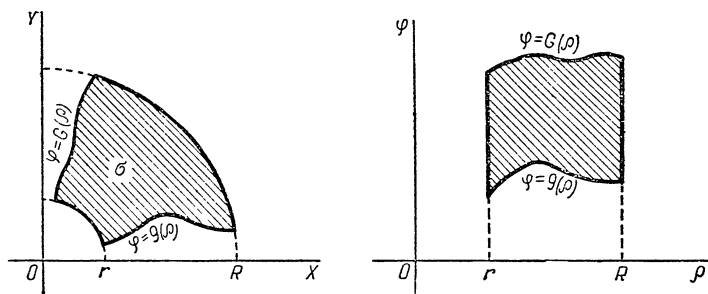


Рис. 396.

Построим на плоскости $\rho O\varphi$ соответствующую область ω (рис. 395, б). По теореме 2 § 10 имеем:

$$\iint_{\sigma} f(x, y) dx dy = \iiint_{\omega} f(\rho \cos \varphi, \rho \sin \varphi) \rho d\rho d\varphi = \int_{\alpha}^{\beta} d\varphi \int_{h(\varphi)}^{H(\varphi)} f(\rho \cos \varphi, \rho \sin \varphi) \rho d\rho.$$

Те же пределы можно получить с помощью чертежа σ , рассуждая примерно следующим образом. У внешнего интеграла пределами являются наименьшее ($\varphi = \alpha$) и наибольшее ($\varphi = \beta$) значения полярного угла φ для точек области σ . Для внутреннего интеграла имеем $\varphi = \text{const}$, $\alpha \leq \varphi \leq \beta$. На плоскости XOY точки, у которых $\varphi = \text{const}$, располагаются на полупрямой, выходящей из начала под углом φ к оси OX . Так как полярная координата ρ точки означает расстояние этой точки от начала O , то нижний и верхний пределы внутреннего интеграла равны соответственно расстоянию от начала O самой близкой и самой далекой точки области σ , лежащей на проведенном луче $\varphi = \text{const}$. Эти расстояния равны $h(\varphi)$ и $H(\varphi)$.

Аналогично в случае области σ , ограниченной линиями, полярные уравнения которых $\varphi = g(\rho)$, $\varphi = G(\rho)$, $\rho = r$, $\rho = R$ ($r < R$, $g(\rho) \leq G(\rho)$, $g(\rho)$ и $G(\rho)$ — непрерывные функции ρ , $r \leq \rho \leq R$) имеем (см. рис. 396):

$$\iint_{\sigma} f(x, y) dx dy = \int_r^R \rho d\rho \int_{g(\rho)}^{G(\rho)} f(\rho \cos \varphi, \rho \sin \varphi) d\varphi.$$

Пример. Вычислить объем тела, ограниченного поверхностями $x^2 + y^2 = 2y$, $z = 0$, $x + y + z = 4$.

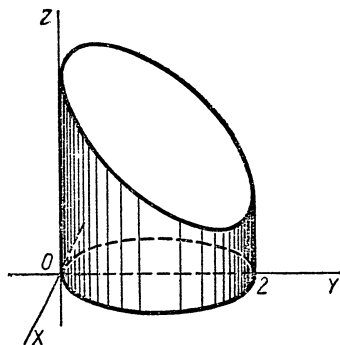


Рис. 397.

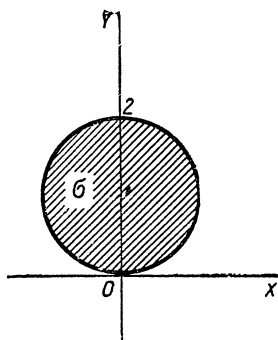


Рис. 398.

Так как данное тело представляет собой цилиндрическое тело, ограниченное сверху поверхностью $z = 4 - x - y$ (рис. 397), то имеем:

$$V = \iint_{\sigma} (4 - x - y) dx dy,$$

где σ — основание тела — круг $x^2 + (y - 1)^2 \leq 1$ в плоскости XOY (см. рис. 398).

Переходя к полярным координатам, получаем:

$$\begin{aligned} V &= \iint_{\sigma} (4 - \rho \cos \varphi - \rho \sin \varphi) \rho d\rho d\varphi = \\ &= \int_0^{\pi} d\varphi \int_0^{2 \sin \varphi} (4 - \rho \cos \varphi - \rho \sin \varphi) \rho d\rho = 3\pi. \end{aligned}$$

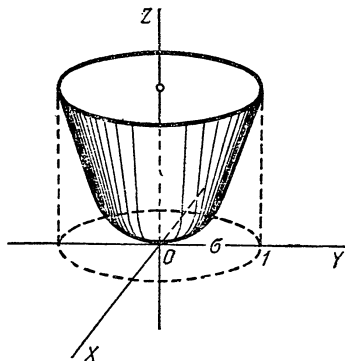


Рис. 399.

Пример. Вычислить площадь части параболоида $x^2 + y^2 = z$, вырезаемой цилиндром $x^2 + y^2 = 1$ (рис. 399).

Так как $\frac{\partial z}{\partial x} = 2x$, $\frac{\partial z}{\partial y} = 2y$, то

$$S = \iint_{\sigma} \sqrt{1 + 4x^2 + 4y^2} dx dy,$$

где σ — круг $x^2 + y^2 \leq 1$ в плоскости XOY .

Переходя к полярным координатам, получаем:

$$S = \iint_{\sigma} \sqrt{1 + 4\rho^2} \rho d\rho d\varphi = \int_0^{2\pi} d\varphi \int_0^1 \rho \sqrt{1 + 4\rho^2} d\rho = \frac{\pi}{6} (5\sqrt{5} - 1).$$

Раздел 4

ТРОЙНЫЕ ИНТЕГРАЛЫ

§ 14. Задача о массе тела

В пространстве XYZ дано тело V , в каждой точке (x, y, z) которого известна плотность распределения массы

$$\mu = \mu(x, y, z)^1).$$

Будем предполагать, что $\mu(x, y, z)$ — непрерывная функция от x, y, z . Требуется вычислить массу m этого неоднородного тела.

¹⁾ Средней плотностью распределения массы в трехмерной области называется отношение массы, сосредоточенной в этой области, к объему области. Плотность распределения массы в данной точке M области есть предел средней плотности распределения массы в области, содержащей эту точку, при стягивании этой области к точке M .

Для решения поставленной задачи разобьем область V на n частей $\Delta V_1, \Delta V_2, \dots, \Delta V_n$, не имеющих общих внутренних точек¹⁾. Выберем в каждой области ΔV_i (внутри или на границе) произвольно точку $M_i(\xi_i, \eta_i, \zeta_i)$. Плотность распределения массы в такой точке равна $\mu(\xi_i, \eta_i, \zeta_i)$. Массу m_i области ΔV_i естественно считать приближенно равной массе однородной области ΔV_i с плотностью распределения $\mu(\xi_i, \eta_i, \zeta_i)$:

$$m_i \approx \mu(\xi_i, \eta_i, \zeta_i) \Delta V_i. \quad (1)$$

Чем мельче область ΔV_i , тем меньше в точках этой области значения функции $\mu(x, y, z)$ отличаются от ее значения в точке M_i , тем, вообще говоря, точнее это приближенное равенство. Суммируя значения масс m_i по всем областям деления, получаем приближенное значение массы m тела V :

$$m \approx \sum_{i=1}^n \mu(\xi_i, \eta_i, \zeta_i) \Delta V_i.$$

Точное значение массы тела V принимается равным пределу, к которому стремится найденное приближенное значение при неограниченном измельчении областей деления и, следовательно, неограниченном увеличении числа областей деления (короче, при стремлении к нулю шага разбиения λ области V).

Таким образом,

$$m = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \mu(\xi_i, \eta_i, \zeta_i) \Delta V_i.$$

Подобного рода пределы часто приходится рассматривать в механике и физике, в связи с чем представляется целесообразным их изучение.

§ 15. Определение тройного интеграла и его свойства

Пусть в замкнутой, кубируемой²⁾ области V пространства XYZ задана произвольная функция $f(x, y, z)$. Разобьем область V на n областей $\Delta V_1, \Delta V_2, \dots, \Delta V_n$, не имеющих общих внутренних точек. В каждой области ΔV_i возьмем произвольно точку $M_i(\xi_i, \eta_i, \zeta_i)$. Значение функции $f(x, y, z)$ в точке M_i умножим на объем ΔV_i i -й области и сложим такие произведения по всем областям деления. Полученная сумма $\sum_{i=1}^n f(\xi_i, \eta_i, \zeta_i) \Delta V_i$ называется *интегральной суммой* для функции $f(x, y, z)$

по области V . Для функции $f(x, y, z)$ можно составить бесчисленное множество интегральных сумм по области V .

Если при стремлении к нулю шага разбиения λ области V существует предел³⁾ интегральных сумм, то этот предел называется *тройным интегралом* от функции $f(x, y, z)$ по области V и обозначается символами:

$$\iiint_V f(x, y, z) dV \quad \text{или} \quad \iiint_V f(x, y, z) dx dy dz.$$

При этом $f(x, y, z)$ — *подинтегральная функция*, V — *область интегрирования*, x, y и z — *переменные интегрирования*, $dV (dx dy dz)$ — *элемент объема*.

¹⁾ Объемы областей будем обозначать теми же буквами, что и сами области.

²⁾ В дальнейшем будут рассматриваться только кубируемые (имеющие конечный объем) области трехмерного пространства.

³⁾ Определение предела дается аналогично соответствующему определению для двумерного случая.

Таким образом,

$$\iiint_V f(x, y, z) dV = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(\xi_i, \eta_i, \zeta_i) \Delta V_i,$$

если этот предел существует.

Из определения следует, что тройной интеграл, так же как и двойной, не зависит от совершаемых при построении интегральных сумм разбиений области V на части и выбора точек на этих частях.

Тройные интегралы являются непосредственным обобщением двойных интегралов на случай трехмерного пространства.

Имеют место следующие достаточные условия существования тройного интеграла.

Теорема 1. Если функция $f(x, y, z)$ непрерывна в замкнутой области V , то тройной интеграл $\iiint_V f(x, y, z) dV$ существует.

Обобщения теоремы 2 (см. § 7, разд. 3) на трехмерный случай в данном курсе не рассматривается.

Решение задачи о массе области V с помощью тройного интеграла запишется в виде:

$$m = \iiint_V \mu(x, y, z) dV,$$

где $\mu(x, y, z)$ — плотность распределения массы.

В этом и заключается *механический смысл* тройного интеграла от неотрицательной функции. Установить геометрический смысл тройного интеграла $\iiint_V f(x, y, z) dV$ для произвольной функции $f(x, y, z)$, не выходя за пределы трехмерного пространства, не представляется возможным.

В случае, если в области V

$$f(x, y, z) \equiv 1,$$

то

$$\iiint_V dV = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n \Delta V_i = V.$$

Таким образом, объем тела V можно вычислять с помощью тройного интеграла по формуле

$$V = \iiint_V dV = \iiint_V dx dy dz.$$

Перечислим простейшие свойства тройных интегралов от функций, непрерывных в рассматриваемых областях.

1. Тройной интеграл $\iiint_V f(x, y, z) dx dy dz$ от обозначения переменных интегрирования не зависит, т. е.

$$\iiint_V f(x, y, z) dx dy dz = \iiint_V f(s, t, u) ds dt du$$

и т. д.

2. Постоянный множитель можно выносить за знак тройного интеграла:

$$\iiint_V k f(x, y, z) dV = k \iiint_V f(x, y, z) dV,$$

где k — число.

3. Тройной интеграл от суммы двух функций равен сумме соответствующих интегралов от слагаемых:

$$\iiint_V [f_1(x, y, z) + f_2(x, y, z)] dV = \iiint_V f_1(x, y, z) dV + \iiint_V f_2(x, y, z) dV.$$

4. Если область V разбита на две области V_1 и V_2 , то

$$\iiint_V f(x, y, z) dV = \iiint_{V_1} f(x, y, z) dV + \iiint_{V_2} f(x, y, z) dV.$$

Примечание. Свойства 3 и 4 верны для любого фиксированного числа слагаемых.

5. Если в области V

$$f(x, y, z) \geq 0,$$

то

$$\iiint_V f(x, y, z) dV \geq 0.$$

6. Если в области V

$$f_1(x, y, z) \leq f_2(x, y, z),$$

то

$$\iiint_V f_1(x, y, z) dV \leq \iiint_V f_2(x, y, z) dV.$$

7.

$$\left| \iiint_V f(x, y, z) dV \right| \leq \iiint_V |f(x, y, z)| dV.$$

8. Теорема о среднем. Если функция $f(x, y, z)$ непрерывна в замкнутой области V , то в этой области существует точка $M(\xi, \eta, \zeta)$, такая, что

$$\iiint_V f(x, y, z) dV = f(\xi, \eta, \zeta) V,$$

где V — объем данной области.

§ 16. Вычисление тройных интегралов

Вычисление тройных интегралов производится путем последовательного вычисления интегралов меньшей кратности.

Теорема 1. Пусть область V ограничена снизу и сверху поверхностями $z = h(x, y)$ и $z = H(x, y)$, где $h(x, y)$ и $H(x, y)$ — непрерывные функции в замкнутой области σ плоскости $ХОУ$, и цилиндрической поверхностью, у которой образующие параллельны оси Oz , а направляющей является граница области σ (рис. 400). Тогда для любой функции $f(x, y, z)$, непрерывной в замкнутой области V , имеет место формула

$$\begin{aligned} & \iiint_V f(x, y, z) dx dy dz = \\ & = \iint_{\sigma} \left\{ \int_{h(x, y)}^{H(x, y)} f(x, y, z) dz \right\} dx dy, \quad (1) \end{aligned}$$

позволяющая свести вычисление тройного интеграла к вычислению двойного интеграла от определенного интеграла (короче, к вычислению повторного интеграла)¹⁾.

Интеграл, стоящий в правой части равенства, обычно записывают в виде:

$$\iint_{\sigma} dx dy \int_{h(x, y)}^{H(x, y)} f(x, y, z) dz.$$

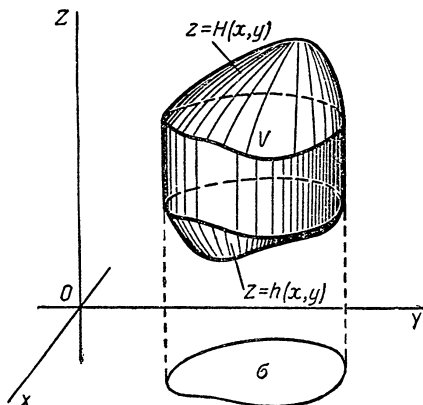


Рис. 400.

¹⁾ Доказательство этой теоремы читатель может получить как непосредственное обобщение на трехмерный случай доказательства теоремы о переходе от двойного интеграла к повторному (см. § 10).

При вычислении тройного интеграла по формуле (1) с помощью повторного интеграла сначала вычисляется внутренний интеграл по переменной z при постоянных x и y (x и y — параметры) в пределах изменения z (для области V) при постоянных x и y , а затем полученная функция x и y интегрируется по переменным x и y по области σ .

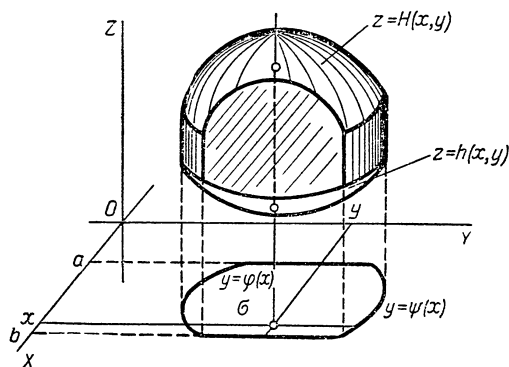


Рис. 401.

Если при этом область σ плоскости XOY ограничена линиями $x=a$, $x=b$ ($a < b$), $y=\phi(x)$, $y=\psi(x)$ [$\phi(x)$ и $\psi(x)$ непрерывны на отрезке $[a, b]$ функции, причем $\phi(x) \leq \psi(x)$] (рис. 401), то, переходя от двойного интеграла по области σ к повторному, получаем формулу

$$\iiint_V f(x, y, z) dx dy dz = \int_a^b dx \int_{\phi(x)}^{\psi(x)} dy \int_{h(x, y)}^{H(x, y)} f(x, y, z) dz, \quad (2)$$

позволяющую вычисление тройного интеграла заменить последовательным вычислением трех определенных интегралов.

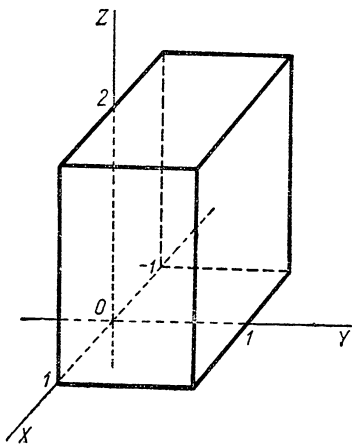


Рис. 402.

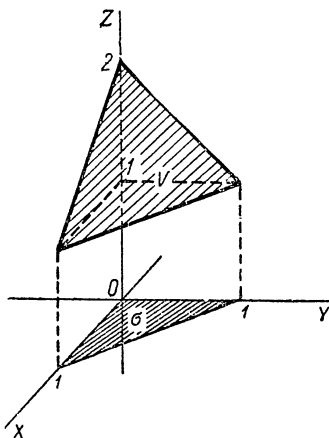


Рис. 403.

В частности, если область V — параллелепипед с гранями $x=a$, $x=b$ ($a < b$), $y=c$, $y=d$ ($c < d$), $z=l$, $z=k$ ($l < k$), то по формуле (2) имеем:

$$\iiint_V f(x, y, z) dx dy dz = \int_a^b dx \int_c^d dy \int_l^k f(x, y, z) dz.$$

Для функции $f(x, y, z)$, равной произведению функций, каждая из которых зависит от одного лишь переменного:

$$f(x, y, z) = f_1(x) \cdot f_2(y) \cdot f_3(z),$$

тройной интеграл по этому параллелепипеду V равен произведению трех определенных интегралов:

$$\iiint_V f(x, y, z) dx dy dz = \int_a^b f_1(x) dx \int_c^d f_2(y) dy \int_l^k f_3(z) dz.$$

Это равенство непосредственно следует из свойств тройных и двойных интегралов.

Пример 1. Вычислить интеграл $\iiint_V (x + y - z) dx dy dz$ по параллелепипеду,

ограниченному плоскостями $x = -1$, $x = +1$, $y = 0$, $y = 1$, $z = 0$, $z = 2$ (рис. 402).

По формуле (2) имеем:

$$\begin{aligned} \iiint_V (x + y - z) dx dy dz &= \\ &= \int_{-1}^1 dx \int_0^1 dy \int_0^2 (x + y - z) dz = \\ &= \int_{-1}^1 dx \int_0^1 (2x + 2y - 2) dy = \\ &= \int_{-1}^1 (2x - 1) dx = -2. \end{aligned}$$

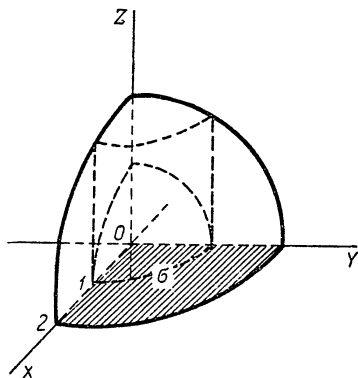


Рис. 404.

Пример 2. Вычислить интеграл $\iiint_V x dx dy dz$ по области V , ограниченной плоскостями $x + y + z = 2$, $z = 1$, $x = 0$, $y = 0$ (рис. 403).

Область V проектируется на плоскость XOY в треугольник σ , ограниченный прямыми $x = 0$, $y = 0$, $x + y = 1$. Применяя формулы 1 и 2, получаем:

$$\begin{aligned} \iiint_V x dx dy dz &= \iint_{\sigma} x dx dy \int_1^{2-x-y} dz = \int_0^1 x dx \int_0^{1-x} dy \int_1^{2-x-y} dz = \\ &= \int_0^1 x dx \int_0^{1-x} (1 - x - y) dy = \int_0^1 \left(\frac{1}{2} x^3 - x^2 + \frac{1}{2} x \right) dx = \frac{1}{24}. \end{aligned}$$

Наряду с указанными выше формулами при соответствующих условиях имеют место формулы, получающиеся из них перестановкой переменных x , y и z .

В случае, если область V может быть разбита на части, удовлетворяющие условиям теоремы 1 или теорем, получающихся из нее перестановкой переменных x , y и z , вычисление тройного интеграла сводится к вычислению соответствующих интегралов по каждой из этих частей.

Пример 3. Составить формулу для вычисления тройного интеграла

$$\iiint_V f(x, y, z) dx dy dz$$

по области V , лежащей в первом октанте и ограниченной поверхностями $x = 0$, $y = 0$, $z = 0$, $x^2 + y^2 + z^2 = 1$, $x^2 + y^2 + z^2 = 4$ (рис. 404). Так как поверхность, ограничивающая область V снизу, задана двумя уравнениями (уравнение плоскости $z = 0$, уравнение сферы $x^2 + y^2 + z^2 = 1$), разбиваем область V на две части цилиндрической поверхностью $x^2 + y^2 = 1$.

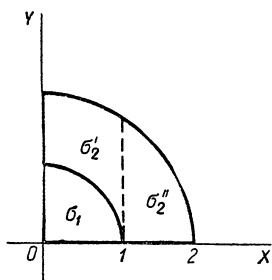


Рис. 405.

Обозначим через σ_1 часть круга $x^2 + y^2 \leq 1$, а через σ_2 — часть кольца $1 \leq x^2 + y^2 \leq 4$, лежащие в первой четверти плоскости XOY , а через V_1 и V_2 — части области V , проектирующиеся соответственно в σ_1 и σ_2 . Тогда

$$\iiint_V f(x, y, z) dx dy dz = \iiint_{V_1} f(x, y, z) dx dy dz + \iiint_{V_2} f(x, y, z) dx dy dz.$$

Интегралы по областям V_1 и V_2 вычисляем, применяя формулы (1) и (2):

$$\begin{aligned} \iiint_{V_1} f(x, y, z) dx dy dz &= \int_{\sigma_1} dx dy \int_0^{\sqrt{4-x^2-y^2}} f(x, y, z) dz = \\ &= \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy \int_0^{\sqrt{4-x^2-y^2}} f(x, y, z) dz, \\ \iiint_{V_2} f(x, y, z) dx dy dz &= \int_{\sigma_2} dx dy \int_0^{\sqrt{4-x^2-y^2}} f(x, y, z) dz = \\ &= \int_0^1 dx \int_{\sqrt{1-x^2}}^2 dy \int_0^{\sqrt{4-x^2-y^2}} f(x, y, z) dz + \int_1^2 dx \int_0^{\sqrt{4-x^2}} dy \int_0^{\sqrt{4-x^2-y^2}} f(x, y, z) dz. \end{aligned}$$

Примечание. Для перехода от двойного интеграла по области σ_2 к повторному надо предварительно область σ_2 разбить на части σ_2' и σ_2'' (см. рис. 405).

§ 17. Некоторые механические приложения тройных интегралов

К вычислению тройных интегралов приводят задачи, связанные с непрерывным распределением массы в пространственной области.

Если известна плотность распределения массы $\mu = \mu(x, y, z)$ области V пространства XYZ ¹⁾, то рассуждая так же, как в двумерном случае (§ 11), получим определения и формулы для вычисления статических моментов области V относительно координатных плоскостей:

$$S_{xy} = \iiint_V z\mu(x, y, z) dV,$$

$$S_{yz} = \iiint_V x\mu(x, y, z) dV,$$

$$S_{zx} = \iiint_V y\mu(x, y, z) dV,$$

моментов инерции относительно координатных осей:

$$J_x = \iiint_V (y^2 + z^2)\mu(x, y, z) dV,$$

$$J_y = \iiint_V (x^2 + z^2)\mu(x, y, z) dV,$$

$$J_z = \iiint_V (x^2 + y^2)\mu(x, y, z) dV,$$

координат центра тяжести:

$$x_c = \frac{\iiint_V x\mu(x, y, z) dV}{\iiint_V \mu(x, y, z) dV} = \frac{S_{yz}}{m},$$

$$y_c = \frac{\iiint_V y\mu(x, y, z) dV}{\iiint_V \mu(x, y, z) dV} = \frac{S_{zx}}{m},$$

$$z_c = \frac{\iiint_V z\mu(x, y, z) dV}{\iiint_V \mu(x, y, z) dV} = \frac{S_{xy}}{m}.$$

¹⁾ Предполагается, что $\mu(x, y, z)$ — непрерывная функция в замкнутой области V .

§ 18. Замена переменных в тройном интеграле

Для тройного интеграла имеет место следующее правило замены переменных.

Если функция $f(x, y, z)$ непрерывна в замкнутой области V , а функции

$$x = \varphi(u, v, w), \quad y = \psi(u, v, w), \quad z = \chi(u, v, w) \quad (1)$$

имеют непрерывные частные производные в замкнутой области T пространства UVW и взаимно однозначно отображают эту область на область V пространства XYZ , то имеет место следующая формула:

$$\iiint_V f(x, y, z) dx dy dz = \iiint_T f(\varphi(u, v, w), \psi(u, v, w), \chi(u, v, w)) |J(u, v, w)| du dv dw, \quad (2)$$

$$\text{где } J(u, v, w) = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix} \text{ — якобиан отображения (1).}$$

Подобно тому как в случае двух переменных модуль якобиана отображения равнялся коэффициенту изменения бесконечно малой площади, модуль якобиана отображения (1) равен коэффициенту изменения бесконечно малого объема при отображении (1).

Примечание. Формула (2) остается справедливой также и в том случае, если отображение области T на область V взаимно однозначно лишь для внутренних точек этих областей.

Остановимся подробнее на двух наиболее употребительных случаях замены переменных в тройных интегралах.

1. Переход в тройном интеграле к цилиндрическим координатам. Положение точки $M(x, y, z)$ пространства XYZ однозначно определяется заданием трех чисел ρ, φ, z , где ρ — длина радиуса-вектора проекции точки M на плоскость XOY , φ —

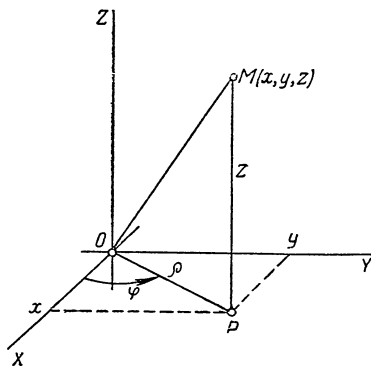


Рис. 406.

угол, составляемый этим радиусом-вектором с осью OX , z — аппликата точки M (рис. 406). Числа ρ, φ, z называются *цилиндрическими координатами* точки M . Они связаны с декартовыми координатами точки M соотношениями:

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi, \quad z = z. \quad (3)$$

По определению $0 \leq \rho < +\infty$, $-\infty < z < +\infty$. Значение угла φ будем брать в пределах $0 \leq \varphi \leq 2\pi$.

На формулы (3), где переменные ρ, φ, z изменяются в указанных пределах, можно смотреть как на задание отображения полосы $\{0 \leq \rho < +\infty, 0 \leq \varphi \leq 2\pi, -\infty < z < +\infty\}$ пространства с прямоугольной системой координат $\rho\varphi z$ (рис. 407) на все пространство XYZ . Отображение (3) взаимно однозначно всюду, кроме границы этой полосы.

При этом отображении (3) плоскости $\rho = \rho_0$ пространства $\rho\varphi z$ соответствуют в пространстве XYZ цилиндрическая поверхность $x^2 + y^2 = \rho_0^2$, а плоскости $\varphi = \varphi_0$ и $z = z_0$ пространства $\rho\varphi z$ отображаются в простран-

стве XYZ соответственно в полуплоскость, проходящую через ось OZ под углом φ_0 к оси OX и плоскость $z=z_0$ (рис. 408).

Якобиан отображения:

$$J(\rho, \varphi, z) = \frac{\partial(x, y, z)}{\partial(\rho, \varphi, z)} = \begin{vmatrix} \cos \varphi & -\rho \sin \varphi & 0 \\ \sin \varphi & \rho \cos \varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho$$

неотрицателен в данной полосе ($\rho=0$ только на границе полосы).

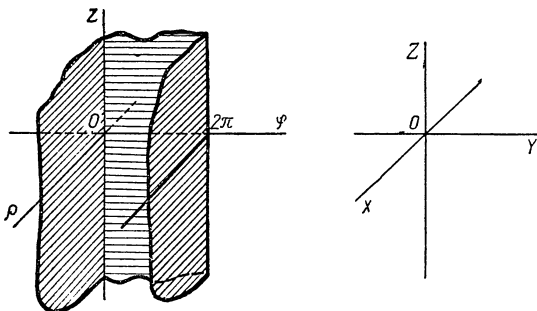


Рис. 407.

Применяя формулу замены переменных в тройном интеграле к отображению (3), получим так называемую *формулу перехода в тройном интеграле от декартовых к цилиндрическим координатам*:

$$\iiint_V f(x, y, z) dx dy dz = \iiint_T f(\rho \cos \varphi, \rho \sin \varphi, z) \rho d\rho d\varphi dz.$$

Здесь T — область пространства $\rho\varphi z$, образом которой при отображении (3) является область V .

Пример. Вычислить объем тела V , ограниченного поверхностями $x^2 + y^2 = z$, $z=1$, в цилиндрических координатах.

Обозначим через T область пространства $\rho\varphi z$, ограниченную поверхностями $\rho^2 = z$, $z=1$, $\varphi=0$, $\varphi=2\pi$ (рис. 409). При отображении (3) ее образом в пространстве XYZ будет данная область V . Поэтому

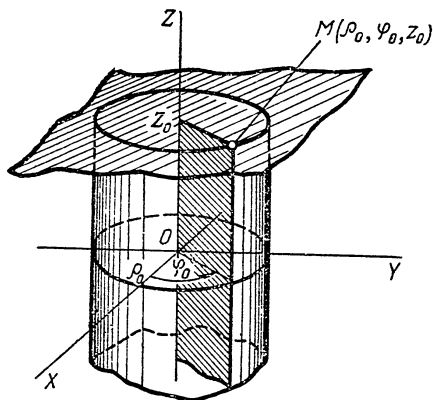


Рис. 408.

$$\begin{aligned} V &= \iiint_V dx dy dz = \\ &= \int_0^1 \int_0^{2\pi} \int_0^1 \rho d\rho d\varphi dz = \int_0^1 d\varphi \int_0^{2\pi} \rho d\rho \int_0^1 dz = \frac{\pi}{2}. \end{aligned}$$

Примечание. Расстановку пределов при вычислении тройного интеграла в цилиндрических координатах можно производить, используя чертеж области V в пространстве XYZ и геометрический смысл цилиндрических координат.

Так, например, при расстановке пределов в полученном выше интеграле можно было рассуждать следующим образом. Так как область V на плоскость XOY проектируется в круг $x^2 + y^2 = 1$, то цилиндрическая координата φ точек тела изменяется в пределах от 0 до 2π . В сечении тела полуплоскостью $\varphi = \text{const}$, $0 \leq \varphi \leq 2\pi$, получаем криволинейный треугольник (см. заштрихованную область на рис. 410). Для его точек цилиндрическая координата ρ изменяется в пределах от $\rho=0$ до $\rho=1$. Постоянному значению ρ , $0 \leq \rho \leq 1$ в пространстве XYZ соответствует цилиндр

$x^2 + y^2 = \rho^2$. В пересечении этого цилиндра с заштрихованной в полуплоскости $\varphi = \text{const}$ областью получаем отрезок прямой, вдоль которого переменная z изменяется от значения z для точки, лежащей на параболоиде, до значения $z=1$ (для точки, лежащей на плоскости $z=1$).

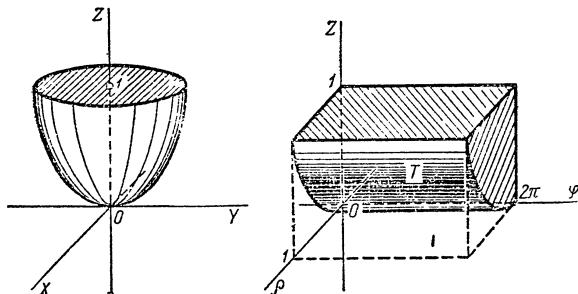


Рис. 409.

Выражая эти значения z в цилиндрических координатах ($z = \rho^2$, $z = 1$), окончательно получаем:

$$V = \int_0^{2\pi} d\varphi \int_0^1 \rho d\rho \int_{\rho^2}^1 dz = \frac{\pi}{2}.$$

2. Переход в тройном интеграле от декартовых к сферическим координатам. Сферическими координатами точки $M(x, y, z)$ пространства XYZ называются числа ρ , φ , θ , где ρ — длина радиуса-вектора $\overline{OM}$ точки M , φ — угол, составляемый с осью OX проекцией этого радиуса-вектора на плоскость XOY , а θ — угол отклонения радиуса-вектора точки M от

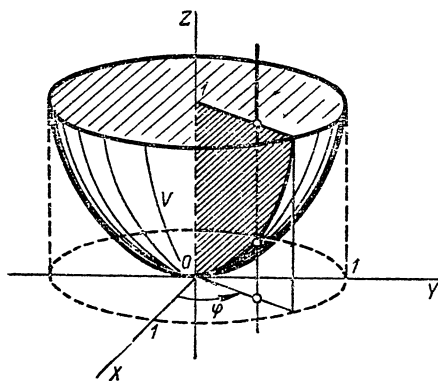


Рис. 410.

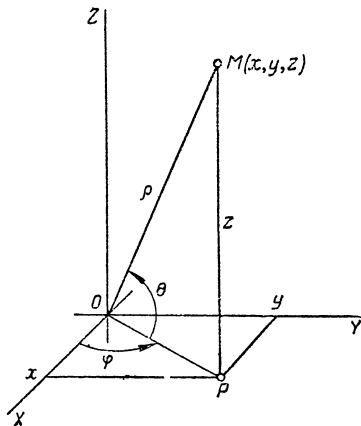


Рис. 411.

плоскости XOY (рис. 411). Задание этих трех чисел однозначно определяет положение точки M в пространстве XYZ .

Сферические координаты ρ , φ , θ точки M связаны с ее декартовыми координатами x , y , z формулами:

$$x = \rho \cos \theta \cos \varphi, \quad y = \rho \cos \theta \sin \varphi, \quad z = \rho \sin \theta. \quad (4)$$

По определению $0 \leq \rho < +\infty$. Значения углов φ и θ будем брать в пределах: $0 \leq \varphi \leq 2\pi$, $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$.

Формулы (4), где переменные ρ , φ , θ изменяются в указанных пределах, можно рассматривать как задание отображения полосы

$$\{0 \leq \rho < +\infty, \quad 0 \leq \varphi \leq 2\pi, \quad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}\}$$

пространства с прямоугольными координатами ρ , φ , θ (рис. 412) на все пространство XYZ .

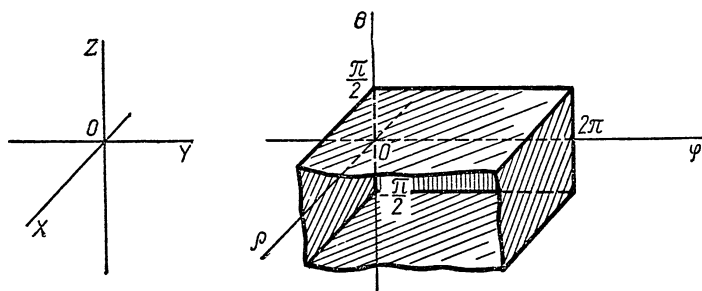


Рис. 412.

При этом отображении плоскости $\rho = \rho_0$ пространства $\rho\varphi\theta$ соответствует сфера $x^2 + y^2 + z^2 = \rho_0^2$ в пространстве XYZ , плоскости $\varphi = \varphi_0$ — полуплоскость, проходящая через ось OZ под углом φ_0 к оси OX , а плоскости $\theta = \theta_0$ — круговой конус с вершиной O , образующие кото-

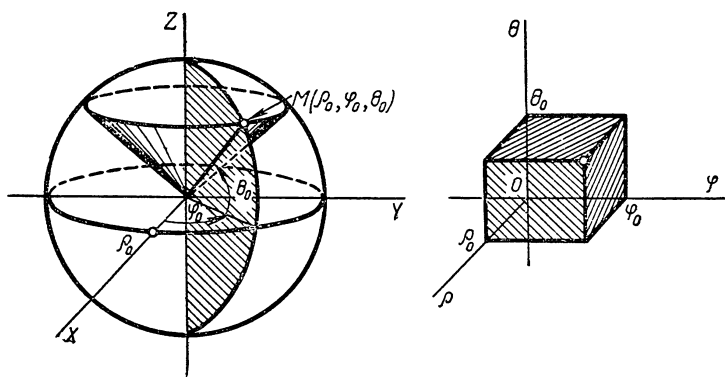


Рис. 413.

рого наклонены к плоскости XOY под углом θ_0 (рис. 413). Это отображение взаимно однозначно всюду, кроме границы указанной выше полосы. Якобиан отображения:

$$\begin{aligned} J(\rho, \varphi, \theta) &= \frac{\partial(x, y, z)}{\partial(\rho, \varphi, \theta)} = \begin{vmatrix} \cos \theta \cos \varphi & -\rho \cos \theta \sin \varphi & -\rho \sin \theta \cos \varphi \\ \cos \theta \sin \varphi & \rho \cos \theta \cos \varphi & -\rho \sin \theta \sin \varphi \\ \sin \theta & 0 & \rho \cos \theta \end{vmatrix} = \\ &= \sin \theta \begin{vmatrix} -\rho \cos \theta \sin \varphi & -\rho \sin \theta \cos \varphi \\ \rho \cos \theta \cos \varphi & -\rho \sin \theta \sin \varphi \end{vmatrix} + \rho \cos \theta \begin{vmatrix} \cos \theta \cos \varphi & -\rho \cos \theta \sin \varphi \\ \cos \theta \sin \varphi & \rho \cos \theta \cos \varphi \end{vmatrix} = \\ &= \rho^2 \cos \theta \sin^2 \theta + \rho^2 \cos^3 \theta = \rho^2 \cos \theta. \end{aligned}$$

В данной полосе якобиан отображения (4) неотрицателен и обращается в нуль лишь на границе полосы.

Применяя формулу замены переменных в тройном интеграле к отображению (4), получим:

$$\iiint_V f(x, y, z) dx dy dz = \iiint_T f(\rho \cos \theta \cos \varphi, \rho \cos \theta \sin \varphi, \rho \sin \theta) \rho^2 \cos \theta d\rho d\varphi d\theta, \quad (5)$$

где T — область пространства $\rho\varphi\theta$, образом которой является область V .

Эта формула называется *формулой перехода в тройном интеграле от декартовых к сферическим координатам*.

При вычислении тройного интеграла в сферических координатах можно пользоваться чертежом области V в пространстве XYZ и геометрическим смыслом сферических координат.

Пример. Вычислить массу полушара V радиуса R , если плотность распределения массы в каждой его точке пропорциональна расстоянию точки от некоторой фиксированной точки O на границе основания полушара.

Выберем начало координат в точке O , а плоскость XOY в плоскости полушара так, чтобы центр шара лежал на оси OY (рис. 414).

Тогда уравнение поверхности, ограничивающей тело V сверху, запишется в виде:

$$x^2 + y^2 + z^2 = 2Ry,$$

плотность распределения массы определяется формулой

$$\mu = k \sqrt{x^2 + y^2 + z^2},$$

а значит, определение массы полушара сведется к вычислению интеграла

$$m = k \iiint_V \sqrt{x^2 + y^2 + z^2} dx dy dz.$$

Применяя формулу (5), получим:

$$m = k \int_0^\pi \int_0^{2\pi} \int_0^{2R \cos \theta \sin \varphi} \rho \cdot \rho^2 \cos \theta d\rho d\varphi d\theta = k \int_0^\pi d\varphi \int_0^{\frac{\pi}{2}} \cos \theta d\theta \int_0^{2R \cos \theta \sin \varphi} \rho^3 d\rho = \frac{4k\pi R^4}{5}.$$

Раздел 5

ПОВЕРХНОСТНЫЕ ИНТЕГРАЛЫ ПЕРВОГО РОДА (ПО ПЛОЩАДИ ПОВЕРХНОСТИ)

§ 19. Определение поверхностного интеграла первого рода

Поверхностный интеграл первого рода представляет собой такое же обобщение двойного интеграла, каким криволинейный интеграл первого рода является по отношению к определенному интегралу.

Пусть S — квадратуемая¹⁾ поверхность в пространстве XYZ , а $F(x, y, z)$ — функция, определенная в точках этой поверхности. Разобьем поверхность S произвольно на n частей $\Delta S_1, \Delta S_2, \dots, \Delta S_n$, не имеющих общих внутренних точек. На каждой поверхности ΔS_i выберем точку

¹⁾ В данном разделе всюду будут рассматриваться только квадратуемые т. е. имеющие определенную площадь, поверхности и их части.

$M_i(\xi_i, \eta_i, \zeta_i)$ (рис. 415). Значение функции $F(x, y, z)$ в точке M_i умножим на площадь ΔS_i поверхности ΔS_i :

$$F(\xi_i, \eta_i, \zeta_i) \Delta S_i$$

и сложим все полученные произведения. Сумма

$$\sum_{i=1}^n F(\xi_i, \eta_i, \zeta_i) \Delta S_i$$

называется *интегральной суммой* для функции $F(x, y, z)$ по поверхности S . Для данной функции $F(x, y, z)$ можно составить бесчисленное множество интегральных сумм по поверхности S .

Если при стремлении к нулю шага λ^1 поверхности существует предел интегральных сумм, то этот предел называется *поверхностным интегралом первого рода* от функции $F(x, y, z)$ по поверхности S и обозначается символом

$$\iint_S F(x, y, z) dS.$$

Таким образом,

$$\begin{aligned} \iint_S F(x, y, z) dS &= \\ &= \lim_{\lambda \rightarrow 0} \sum_{i=1}^n F(\xi_i, \eta_i, \zeta_i) \Delta S_i, \end{aligned}$$

если этот предел существует.

Из определения предела следует, что интеграл $\iint_S F(x, y, z) dS$ не зависит от способа разбиения поверхности S на части и от выбора точек M_i на этих частях.

Примечание. Поверхностные интегралы первого рода часто называются *поверхностными интегралами по площади* поверхности S в отличие от *поверхностных интегралов второго рода* (по координатам), которые будут рассматриваться ниже.

Легко видеть, что поверхностные интегралы первого рода обладают следующими простейшими свойствами:

1. $\iint_S dS = S$.
2. $\iint_S kF(x, y, z) dS = k \iint_S F(x, y, z) dS$, где k — число.
3. $\iint_S [F_1(x, y, z) + F_2(x, y, z)] dS = \iint_S F_1(x, y, z) dS + \iint_S F_2(x, y, z) dS$.
4. Если поверхность S разбита на части S_1 и S_2 , то

$$\iint_S F(x, y, z) dS = \iint_{S_1} F(x, y, z) dS + \iint_{S_2} F(x, y, z) dS.$$
5. Если $F_1(x, y, z) \leq F_2(x, y, z)$, то

$$\iint_S F_1(x, y, z) dS \leq \iint_S F_2(x, y, z) dS.$$
6.

$$\left| \iint_S F(x, y, z) dS \right| \leq \iint_S |F(x, y, z)| dS.$$

¹⁾ *Диаметром* поверхности называется наибольшее из расстояний между точками поверхности. Шаг разбиения поверхности на части — наибольший из диаметров поверхностей деления.

7. Теорема о среднем. Если функция $F(x, y, z)$ непрерывна на поверхности $S^1)$, то на этой поверхности существует точка (ξ, η, ζ) , такая, что

$$\iint_S F(x, y, z) dS = F(\xi, \eta, \zeta) \cdot S,$$

где S — площадь поверхности.

При этом предполагается, что каждый из рассматриваемых в свойствах 1—6 интегралов существует.

§ 20. Условия существования и вычисление поверхностных интегралов первого рода

Приведем простейшие достаточные условия существования поверхностного интеграла первого рода.

Теорема 1. Если поверхность S может быть задана уравнением $z = f(x, y)$, где функция $f(x, y)$ и ее частные производные $f'_x(x, y)$ и $f'_y(x, y)$ непрерывны в замкнутой ограниченной области σ — проекции S на плоскость XOY , а функция $F(x, y, z)$ непрерывна на поверхности S , то интеграл

$$\iint_S F(x, y, z) dS$$

существует.

К использованию этих условий, равно как и условий, получающихся из них перестановкой переменных x, y и z , сводится большинство практически встречающихся случаев.

Вычисление поверхностных интегралов первого рода обычно производится путем их сведения к двойным интегралам.

Пусть функция $F(x, y, z)$ непрерывна на поверхности S , удовлетворяющей условиям теоремы 1. Тогда для любого способа разбиения поверхности S на части и любого выбора точек M_i на этих частях

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n F(\xi_i, \eta_i, \zeta_i) \Delta S_i = \iint_S F(x, y, z) dS.$$

Построим интегральные суммы специальным образом. А именно, разобьем область σ плоскости XOY на части $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$, не имеющие общих внутренних точек. Обозначим через ΔS_i часть поверхности S , проектирующуюся на плоскость XOY в область $\Delta\sigma_i$ (рис. 416). При этом поверхность S окажется разбитой на n частей: $\Delta S_1, \Delta S_2, \dots, \Delta S_n$.

Площадь каждой части ΔS_i может быть найдена по формуле (см. § 11)

$$\Delta S_i = \iint_{\Delta\sigma_i} \sqrt{1 + f'^2_x(x, y) + f'^2_y(x, y)} dx dy.$$

¹⁾ При этом функция $F(x, y, z)$ называется *непрерывной на поверхности S в точке M* , если для любой последовательности точек $M_1, M_2, \dots, M_n, \dots$ поверхности S , сходящейся к точке M , выполнено равенство: $\lim_{n \rightarrow \infty} F(M_n) = F(M)$. Функция, непрерывная в каждой точке поверхности S , называется *непрерывной на поверхности S* .

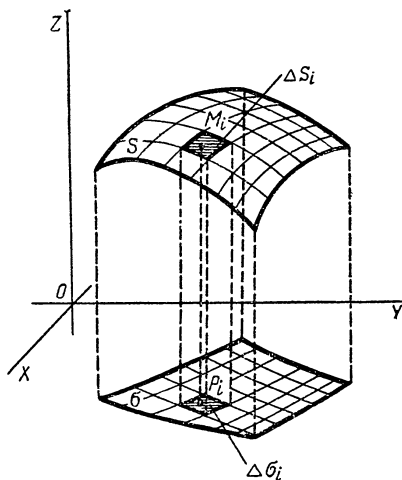


Рис. 416.

Применяя к двойному интегралу теорему о среднем, получим:

$$\Delta S_i = \sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)} \Delta \sigma_i,$$

где $P_i(\xi_i, \eta_i)$ — некоторая точка области $\Delta \sigma_i$.

Обозначим через M_i точку поверхности ΔS_i с координатами (ξ_i, η_i, ζ_i) , $\zeta_i = f(\xi_i, \eta_i)$, и составим интегральную сумму для интеграла от функции $F(x, y, z)$ по поверхности S , соответствующую произведенному разбиению поверхности S на части и выбранным точкам M_i , $i = 1, 2, \dots, n$, на этих частях.

По доказанному

$$\sum_{i=1}^n F(\xi_i, \eta_i, \zeta_i) \Delta S_i = \sum_{i=1}^n F(\xi_i, \eta_i, f(\xi_i, \eta_i)) \sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)} \Delta \sigma_i.$$

В правой части равенства имеем интегральную сумму для двойного интеграла от функции $F(x, y, f(x, y)) \cdot \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)}$,

непрерывной в области σ .

При стремлении к нулю шага разбиения области σ шаг разбиения поверхности S (в силу непрерывности функций $f_x'(x, y)$ и $f_y'(x, y)$) также стремится к нулю. В пределе получаем формулу

$$\iint_S F(x, y, z) dS = \iint_{\sigma} F(x, y, f(x, y)) \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)} dx dy, \quad (1)$$

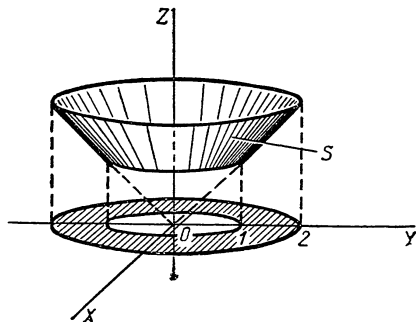


Рис. 417.

выражающую интеграл по поверхности S через двойной интеграл по проекции поверхности S на плоскость XOY .

Примечание. Можно доказать, что интеграл по поверхности S существует и формула (1) имеет место также и в том случае, когда функции $f_x'(x, y)$, $f_y'(x, y)$ ограничены в области σ и непрерывны всюду, кроме отдельных кусочно-гладких линий, а функция $F(x, y, z)$ ограничена на поверхности S и непрерывна на ней всюду, кроме отдельных кусочно-гладких линий.

Аналогично получаются формулы, выражающие при соответствующих условиях интеграл по поверхности S через двойные интегралы по ее проекции на плоскости ZOX и YOZ .

При вычислении интегралов по поверхностям более сложного вида применяется предварительное разбиение этих поверхностей на части указанного вида.

Пример. Вычислим интеграл $\iint_S z dS$ по части поверхности конуса $z = \sqrt{x^2 + y^2}$, $1 \leq z \leq 2$ (рис. 417).

Поверхность S проектируется на плоскость XOY в область σ , представляющую собой кольцо $1 \leq x^2 + y^2 \leq 4$. В этом кольце функции $f(x, y) = \sqrt{x^2 + y^2}$, $f_x'(x, y) = \frac{x}{\sqrt{x^2 + y^2}}$, $f_y'(x, y) = \frac{y}{\sqrt{x^2 + y^2}}$ — непрерывны. Следовательно,

$$\iint_S z dS = \iint_{\sigma} \sqrt{x^2 + y^2} \sqrt{2} dx dy = \sqrt{2} \int_0^{2\pi} d\varphi \int_1^2 \rho^2 d\rho = 2\pi \sqrt{2} \cdot \frac{7}{3}.$$

§ 21. Некоторые приложения поверхностных интегралов первого рода

С помощью поверхностных интегралов первого рода можно определять массы, статические моменты, моменты инерции, координаты центров тяжести и тому подобные величины для материальных поверхностей с известной плотностью распределения массы¹⁾. Решение этих задач аналогично решению соответствующих задач для случая материальной кривой, материальной плоской и пространственной области.

Рассмотрим подробнее задачу о притяжении материальной точки материальной поверхностью S .

Как известно, по закону Ньютона материальная точка $M(x, y, z)$ массы m притягивает материальную точку $A(a, b, c)$ единичной массы с силой, направленной от A к M и численно равной $\frac{km}{r^2}$, где $r = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$ — расстояние AM , а k коэффициент пропорциональности, зависящей от выбора основных единиц измерения. Так как направляющие косинусы вектора $\overrightarrow{AM}$ равны $\frac{x-a}{r}$, $\frac{y-b}{r}$, $\frac{z-c}{r}$, вектор силы имеет координаты

$$km \frac{x-a}{r^3}, \quad km \frac{y-b}{r^3}, \quad km \frac{z-c}{r^3}.$$

Если точка A притягивается системой материальных точек $M_1, M_2, \dots, M_n$ с массами $m_1, m_2, \dots, m_n$, то результирующая сила получается сложением отдельных сил притяжения. Проекции этой результирующей силы на координатные оси равны алгебраическим суммам проекций слагаемых.

Пусть теперь притягивающая масса m распределена непрерывным образом по поверхности S с плотностью распределения $\mu = \mu(x, y, z)$. Определим силу $F\{P, Q, R\}$, с которой эта материальная поверхность S притягивает единицу массы, помещенную в точку $A(a, b, c)$, лежащую вне поверхности S .

Для решения поставленной задачи разобьем поверхность S на n частей $\Delta S_1, \Delta S_2, \dots, \Delta S_n$, на каждой части ΔS_i выберем произвольную точку $M_i(\xi_i, \eta_i, \zeta_i)$. Если диаметр поверхности ΔS_i мал, то плотность распределения массы в точках этой поверхности мало отличается от ее значения $\mu(\xi_i, \eta_i, \zeta_i)$ в точке M_i . Массу m_i поверхности ΔS_i естественно считать приближенно равной массе однородной поверхности ΔS_i с плотностью распределения $\mu(\xi_i, \eta_i, \zeta_i)$:

$$m_i \approx \mu(\xi_i, \eta_i, \zeta_i) \Delta S_i.$$

Силу, с которой на материальную точку A действует поверхность ΔS_i , естественно считать приближенно равной силе, с которой на точку A действует масса m_i , сосредоточенная в точке M_i . Ее проекции на оси координат соответственно равны:

$$km_i \frac{\xi_i - a}{r_i^3}, \quad km_i \frac{\eta_i - b}{r_i^3}, \quad km_i \frac{\zeta_i - c}{r_i^3},$$

где

$$r_i = \sqrt{(\xi_i - a)^2 + (\eta_i - b)^2 + (\zeta_i - c)^2},$$

или, если заменить массу m_i ее приближенным значением,

$$k\mu(\xi_i, \eta_i, \zeta_i) \frac{\xi_i - a}{r_i^3} \Delta S_i, \quad k\mu(\xi_i, \eta_i, \zeta_i) \frac{\eta_i - b}{r_i^3} \Delta S_i, \quad k\mu(\xi_i, \eta_i, \zeta_i) \frac{\zeta_i - c}{r_i^3} \Delta S_i.$$

¹⁾ Плотность распределения массы в точке M поверхности S определяется аналогично плоскому случаю.

²⁾ Поверхности $\Delta S_1, \Delta S_2, \dots, \Delta S_n$ не должны иметь общих внутренних точек.

При таких предположениях сила F оказывается приближенно равной силе, с которой на точку A действует фиктивная система из n материальных точек $M_1, M_2, \dots, M_n$ с массами $m_1, m_2, \dots, m_n$, в сумме равными массе поверхности S , а значит, проекции P, Q, R силы F на оси координат приближенно равны:

$$\sum_{i=1}^n k_{\mu}(\xi_i, \eta_i, \zeta_i) \frac{\xi_i - a}{r_i^3} \Delta S_i, \quad \sum_{i=1}^n k_{\mu}(\xi_i, \eta_i, \zeta_i) \frac{\eta_i - b}{r_i^3} \Delta S_i, \\ \sum_{i=1}^n k_{\mu}(\xi_i, \eta_i, \zeta_i) \frac{\zeta_i - c}{r_i^3} \Delta S_i.$$

Переходя к пределу при стремлении к нулю диаметров всех поверхностей деления, получаем точные значения проекции силы притяжения F на оси координат в виде поверхностных интегралов первого рода:

$$P = \iint_S k_{\mu}(x, y, z) \frac{x-a}{r^3} dS, \quad Q = \iint_S k_{\mu}(x, y, z) \frac{y-b}{r^3} dS, \\ R = \iint_S k_{\mu}(x, y, z) \frac{z-c}{r^3} dS,$$

где $r = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$.

Раздел 6

ПОВЕРХНОСТНЫЕ ИНТЕГРАЛЫ ВТОРОГО РОДА (ПО КООРДИНАТАМ)

Поверхностные интегралы второго рода можно рассматривать как обобщение двойных интегралов, подобное тому, каким криволинейные интегралы второго рода являются по отношению к определенным интегралам. При изучении криволинейных интегралов второго рода рассматривалась направленная кривая, в соответствующих интегральных суммах значения функции в точках кривой умножались на взятые с определенным знаком длины проекций дуг деления на координатные оси. В двумерном случае понятие направленной кривой заменяется понятием стороны поверхности, а при составлении интегральных сумм значения функции в точках поверхности умножаются на взятые с определенным знаком площади проекций поверхностей деления на координатные плоскости.

§ 22. Двухсторонние поверхности. Ориентация поверхности.

Возьмем на гладкой поверхности S^1) произвольную точку M , проведем в этой точке нормаль к поверхности и выберем одно определенное направление нормали. Начнем теперь непрерывно перемещать точку M по замкнутому контуру, лежащему на поверхности S и не пересекающему ее границы, приписывая каждому из ее положений то из двух возможных направлений нормали, в которое непрерывно переходит исходное положение нормали ^{2}). В прежнее положение точка M вернется либо с тем же направлением нормали, что и исходное, либо с прямо противоположным направлением нормали (рис. 418).

Если на поверхности S есть хотя бы одна точка и хотя бы один не пересекающий границу поверхности контур, при обходе по которому

¹) Определение гладкой поверхности см. в § 11, стр. 533.

²) Это значит, что угол между выбранными направлениями нормалей в исходной и движущей точке контура изменяется непрерывно.

направление нормали в точке меняется на противоположное, то поверхность S называется *односторонней*. Примером такой поверхности может служить лист Мебиуса, модель которого изображена на рис. 419.

Если для любой точки поверхности S и любого замкнутого, не пересекающего границы поверхности S контура после обхода направление нормали не меняется, то поверхность называется *двухсторонней*, а совокупность всех точек поверхности с выбранными в них по непрерывности направлениями нормали называется *стороной* такой поверхности.

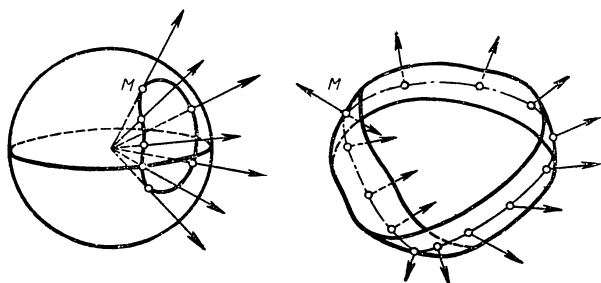


Рис. 418.

Для двухсторонней поверхности выбор направления нормали в одной ее точке определяет направления нормалей во всех точках поверхности, т. е. определяет сторону поверхности.

Примерами двухсторонних поверхностей могут служить любая плоскость, сфера, эллипсоид, параболоид, любая поверхность, задаваемая уравнением $z = f(x, y)$, где $f(x, y)$, $f'_x(x, y)$ и $f'_y(x, y)$ — функции, непрерывные в некоторой области σ плоскости XOY и т. д. Данное выше

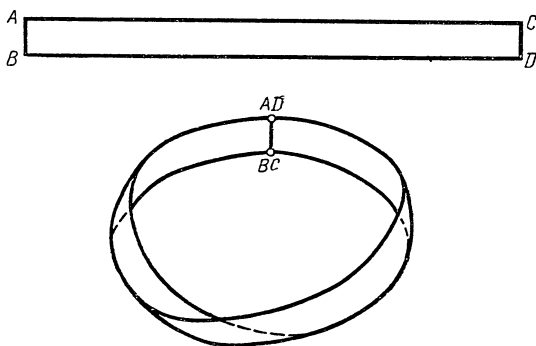


Рис. 419.

определение стороны поверхности полностью согласуется с интуитивным представлением о «верхней» и «нижней» сторонах поверхности $z = f(x, y)$, о «внутренней» и «внешней» сторонах замкнутой поверхности, например сферы, эллипсоида и пр.

Пусть S — незамкнутая двухсторонняя поверхность, ограниченная контуром L , не имеющим точек самопересечения. Выберем определенную сторону поверхности. Будем называть *положительным направлением* обхода контура L то, при движении по которому по выбранной стороне поверхности сама поверхность остается слева (рис. 420).

Аналогично устанавливается положительное направление обхода для любого замкнутого контура (без точек самопересечения), лежащего на поверхности S . Направление обхода контура, противоположное положительному, называется *отрицательным*.

Двухсторонняя поверхность с установленным на ней таким образом положительным направлением обхода контуров называется *ориентированной* поверхностью. Если изменить сторону поверхности, то положительные и отрицательные направления обхода контуров поменяются ро-

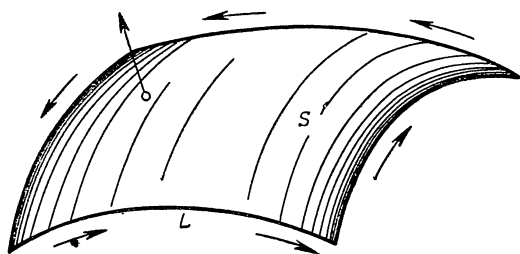


Рис. 420.

лями («поверхность изменит свою ориентацию»). Таким образом, выбор стороны поверхности определяет ее ориентацию и, наоборот, выбор положительного направления обхода контура L , ограничивающего поверхность, однозначно определяет ее сторону.

Рассмотрим теперь кусочно-гладкую поверхность S , т. е. поверхность S , состоящую из любого числа m двух-

сторонних кусков $S_1, S_2, \dots, S_m$ рассмотренного выше вида, соединенных между собой по некоторым «ребрам» (рис. 421). В точках этих ребер касательная плоскость и нормаль к поверхности S может не существовать, т. е. данное выше определение двухсторонней поверхности к поверхности S , вообще говоря, неприменимо. Дополним это определение. А именно, будем называть данную поверх-

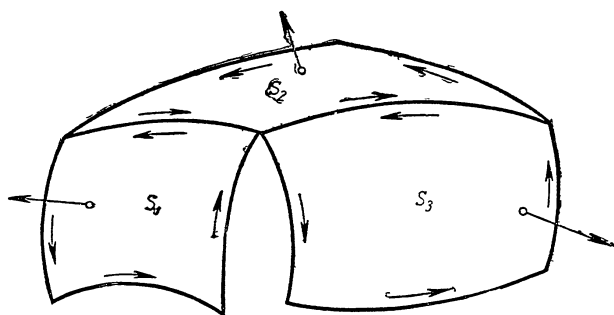


Рис. 421.

ность S *двухсторонней*, если на контурах каждого из кусков $S_1, S_2, \dots, S_m$ можно так задать направления обхода, что общие части этих контуров («ребра») будут описываться во взаимно противоположных направлениях. Выбирая направления обхода контуров $S_1, S_2, \dots, S_m$ указанным образом, мы тем самым на каждом из кусков $S_1, S_2, \dots, S_m$ определим сторону поверхности. Сторона поверхности S определится как совокупность сторон ее частей.

§ 23. Определение поверхностных интегралов второго рода (по координатам) и их простейшие свойства

Рассмотрим в пространстве XYZ двухстороннюю поверхность S , состоящую из конечного числа кусков, каждый из которых либо задан уравнением вида $z = f(x, y)$, либо является цилиндрической поверхностью с образующими, параллельными оси OZ^1). Пусть $R(x, y, z)$ — функция, определенная в точках поверхности S . Выберем определенную сторону поверхности S . Затем разобьем поверхность S на n частей $\Delta S_1, \Delta S_2, \dots, \Delta S_n$,

¹⁾ Предполагается, что эти цилиндрические куски поверхности на плоскость XOY проектируются в кусочно-гладкие кривые.

не имеющих общих внутренних точек и таких, что каждая часть ΔS_i целиком уместается на одном из указанных выше кусков поверхности S .

Обозначим через $(\Delta S_i)_{xy}$ площадь проекции на плоскость XOY поверхности деления ΔS_i , взятую со знаком плюс, если на ΔS_i выбранная сторона поверхности направлена в сторону возрастания z (иначе, если в точках ΔS_i выбранные направления нормалей составляют острые углы с осью OZ), и со знаком минус, если на ΔS_i выбранная сторона поверхности направлена в сторону убывания z (выбранные направления нормалей в точках ΔS_i составляют тупые углы с осью OZ) (рис. 422). Если поверхность деления лежит на цилиндрическом куске поверхности S с образующими, параллельными оси OZ , то проекция ΔS_i на плоскость XOY представляет собой дугу кривой, так что вопрос о знаке $(\Delta S_i)_{xy}$ в этом случае отпадает, $(\Delta S_i)_{xy} = 0$.

Возьмем на каждой части ΔS_i поверхности S произвольно точку $M_i(\xi_i, \eta_i, \zeta_i)$ и умножим значение функции $R(x, y, z)$ в точке M_i на $(\Delta S_i)_{xy}$:

$$R(\xi_i, \eta_i, \zeta_i)(\Delta S_i)_{xy}.$$

Сумма всех таких произведений

$$\sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i)(\Delta S_i)_{xy}$$

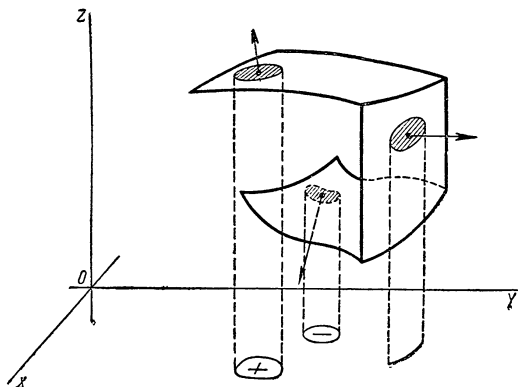


Рис. 422.

называется *интегральной суммой* для функции $R(x, y, z)$ по поверхности S по переменным x и y . Естественно, что таких сумм для заданной на поверхности S функции $R(x, y, z)$ можно составить бесчисленное множество.

Если при стремлении к нулю шага разбиения λ поверхности S интегральные суммы имеют предел¹⁾, то этот предел называют *поверхностным интегралом (второго рода)* по выбранной стороне поверхности S от функции $R(x, y, z)$ по переменным x и y и обозначают символом

$$\iint_S R(x, y, z) dx dy. \quad (1)$$

Так как символ (1) не содержит указания на сторону поверхности S , ее приходится задавать дополнительно. Таким образом,

$$\iint_S R(x, y, z) dx dy = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i)(\Delta S_i)_{xy},$$

если этот предел существует.

Аналогичным образом (при соответствующих условиях на вид поверхности S) определяются поверхностные интегралы второго рода по выбранной стороне поверхности S по переменным y и z (z и x) от функции $P(x, y, z)$ ($Q(x, y, z)$), определенной на поверхности S :

$$\iint_S P(x, y, z) dy dz = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n P(\xi_i, \eta_i, \zeta_i)(\Delta S_i)_{yz}, \quad (2)$$

$$\iint_S Q(x, y, z) dz dx = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n Q(\xi_i, \eta_i, \zeta_i)(\Delta S_i)_{zx}. \quad (3)$$

¹⁾ Определение предела интегральных сумм читатель даст самостоятельно по аналогии с соответствующим определением для двойного интеграла.

Если поверхность S такова, что для функций $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$, определенных в ее точках, существуют интегралы (1), (2) и (3), то можно ввести поверхностный интеграл «общего» вида по выбранной стороне поверхности:

$$\iint_S P(x, y, z) dy dz + Q(x, y, z) dz dx + R(x, y, z) dx dy.$$

Он определяется равенством

$$\begin{aligned} & \iint_S P(x, y, z) dy dz + Q(x, y, z) dz dx + R(x, y, z) dx dy = \\ & = \iint_S P(x, y, z) dy dz + \iint_S Q(x, y, z) dz dx + \iint_S R(x, y, z) dx dy. \end{aligned} \quad (4)$$

К вычислению поверхностных интегралов второго рода приводит, например, решение так называемой **задачи о потоке векторного поля**.

Если в каждой точке $M(x, y, z)$ пространства XYZ (или некоторой его области) задан вектор скорости $\mathbf{a} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$ проходящих через эту точку частиц движущейся жидкости¹⁾, то количество жидкости, протекающее за единицу времени через поверхность S (в выбранную сторону), равно интегралу

$$\iint_S P(x, y, z) dy dz + Q(x, y, z) dz dx + R(x, y, z) dx dy$$

по выбранной стороне поверхности S .

Решение этой задачи в данном курсе не рассматривается, она обычно подробно изучается в теории поля.

Из определения поверхностных интегралов второго рода вытекают следующие их свойства²⁾.

1. Любой поверхностный интеграл (второго рода) изменяет знак при перемене стороны поверхности.

2. Постоянный множитель можно выносить за знак поверхностного интеграла:

$$\begin{aligned} \iint_S k P dy dz &= k \iint_S P dy dz, \\ \iint_S k Q dz dx &= k \iint_S Q dz dx, \\ \iint_S k R dx dy &= k \iint_S R dx dy, \end{aligned}$$

где k — любое число.

3. Поверхностный интеграл от суммы двух функций равен сумме соответствующих интегралов от слагаемых:

$$\begin{aligned} \iint_S (P_1 + P_2) dy dz &= \iint_S P_1 dy dz + \iint_S P_2 dy dz, \\ \iint_S (Q_1 + Q_2) dz dx &= \iint_S Q_1 dz dx + \iint_S Q_2 dz dx, \\ \iint_S (R_1 + R_2) dx dy &= \iint_S R_1 dx dy + \iint_S R_2 dx dy. \end{aligned}$$

4. Если поверхность S разбита на части S_1 и S_2 , то интеграл по всей поверхности S равен сумме интегралов по ее частям:

$$\iint_S = \iint_{S_1} + \iint_{S_2}.$$

¹⁾ Жидкость предполагается несжимаемой (имеющей постоянную плотность) в рассматриваемой области.

²⁾ Каждый из рассматриваемых интегралов предполагается существующим.

5. Если S — цилиндрическая поверхность с образующими, параллельными оси OZ , то

$$\iint_S R(x, y, z) dx dy = 0.$$

6. Если S — цилиндрическая поверхность с образующими, параллельными оси OX , то

$$\iint_S P(x, y, z) dy dz = 0.$$

7. Если S — цилиндрическая поверхность с образующими, параллельными оси OY , то

$$\iint_S Q(x, y, z) dz dx = 0.$$

§ 24. Существование и вычисление поверхностных интегралов второго рода

Пусть S — двухсторонняя поверхность, заданная уравнением $z = f(x, y)$, где $f(x, y)$ — непрерывная функция в замкнутой области σ — проекции поверхности S на плоскость XOY ¹⁾, а $R(x, y, z)$ — непрерывная функция на поверхности S . Выберем верхнюю сторону поверхности S .

Разобьем поверхность S сетью кусочно-гладких линий на n частей $\Delta S_1, \Delta S_2, \dots, \Delta S_n$. Проекциями этих линий на плоскость XOY область σ разобьется на n областей, обозначим их соответственно $\Delta \sigma_1, \Delta \sigma_2, \dots, \Delta \sigma_n$ (рис. 423). Возьмем на каждой части ΔS_i произвольную точку $M_i(\xi_i, \eta_i, \zeta_i)$ и составим интегральную сумму

$$\sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i) (\Delta S_i)_{xy}.$$

Так как сторона поверхности выбрана верхняя, то $(\Delta S_i)_{xy} = = \Delta \sigma_i$ ²⁾, а значит

$$\begin{aligned} \sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i) (\Delta S)_{xy} &= \\ &= \sum_{i=1}^n R(\xi_i, \eta_i, f(\xi_i, \eta_i)) \Delta \sigma_i. \end{aligned}$$

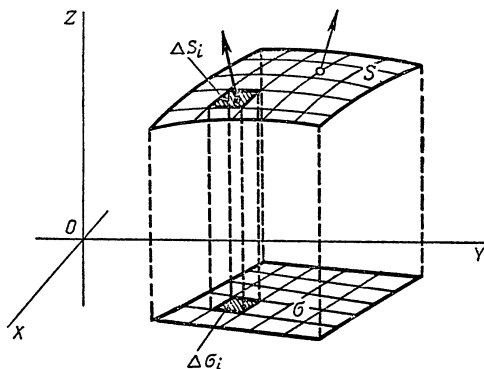


Рис. 423.

В правой части равенства имеем интегральную сумму для двойного интеграла от непрерывной функции $R(x, y, f(x, y))$ по области σ .

При стремлении к нулю шага разбиения поверхности S шаг разбиения области σ также стремится к нулю. В пределе получаем формулу

$$\iint_S R(x, y, z) dx dy = \iint_{\sigma} R(x, y, f(x, y)) dx dy, \quad (1)$$

выражающую интеграл по верхней стороне поверхности S по переменным x и y через двойной интеграл по ее проекции на плоскость XOY .

Примечание. Интегрирование по нижней стороне поверхности S сводится к интегрированию по верхней стороне после перемены знака перед интегралом.

¹⁾ Предполагается, что область σ квадрируема.

²⁾ Здесь, как и всюду в дальнейшем, области и их площади обозначаются одинаково.

Аналогичным образом устанавливается справедливость формул:

$$\iint_S P(x, y, z) dy dz = \iint_{\sigma} P(f(y, z), y, z) dy dz \quad (2)$$

для функции $P(x, y, z)$, непрерывной на двухсторонней поверхности S , заданной уравнением $x=f(y, z)$ (σ — проекция поверхности S на плоскость YOZ , сторона поверхности S направлена в сторону роста x), и

$$\iint_S Q(x, y, z) dz dx = \iint_{\sigma} Q(x, f(x, z), z) dz dx \quad (3)$$

для функции $Q(x, y, z)$, непрерывной на двухсторонней поверхности S , заданной уравнением $y=f(x, z)$ (σ — проекция поверхности S на плоскость ZOX , сторона поверхности S направлена в сторону роста переменной y)¹⁾. Формулы (1), (2), (3) применяются при вычислении соответствующих поверхностных интегралов, а предложенный вывод этих формул дает одновременно доказательство существования рассматриваемых интегралов.

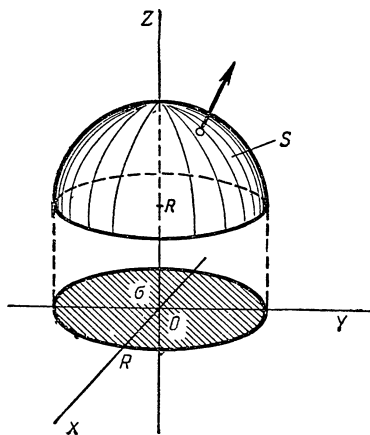


Рис. 424.

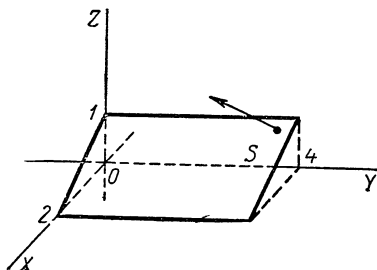


Рис. 425.

Пример 1. Вычислить интеграл $\iint_S (z-R)^2 dx dy$ по верхней стороне полусферы $x^2 + y^2 + z^2 = 2Rz$, $R \leq z \leq 2R$ (рис. 424).

Данную поверхность S можно задать уравнением $z = R + \sqrt{R^2 - x^2 - y^2}$. Поэтому по формуле (1) имеем:

$$\iint_S (z-R)^2 dx dy = \iint_{\sigma} (R^2 - x^2 - y^2) dx dy,$$

где σ — круг $x^2 + y^2 \leq R^2$ плоскости XOY , в который проектируется поверхность S . Вычисляя двойной интеграл, получим:

$$\iint_S (z-R)^2 dx dy = \int_0^{2\pi} d\varphi \int_0^R (R^2 - r^2) r dr = \frac{\pi R^4}{2}.$$

Пример 2. Вычислить интеграл $\iint_S x dy dz + y dz dx + z dx dy$ по верхней стороне части плоскости $x + 2z = 2$, лежащей в первом октанте, и отсекаемой плоскостью $y = 4$ (рис. 425).

По определению

$$\iint_S x dy dz + y dz dx + z dx dy = \iint_S x dy dz + \iint_S y dz dx + \iint_S z dx dy.$$

При этом

$$\iint_S x dy dz = \iint_{\sigma_1} (2 - 2z) dy dz = 2 \int_0^4 dy \int_0^1 (1 - z) dz = 4$$

¹⁾ Функция f предполагается непрерывной в замкнутой квадратуемой области σ .

по формуле (2),

$$\iint_S y \, dz \, dx = 0,$$

так как плоскость S параллельна оси OY ,

$$\iint_S z \, dx \, dy = \int_{\sigma_2} \int_0^2 \left(1 - \frac{x}{2}\right) dx \, dy = \int_0^4 dy \int_0^2 \left(1 - \frac{x}{2}\right) dx = +4$$

по формуле (1).

Следовательно,

$$\iint_S x \, dy \, dz + y \, dz \, dx + z \, dx \, dy = 4 + 0 + 4 = 8.$$

Если двухстороннюю поверхность S можно разбить на конечное число кусков, каждый из которых либо задается уравнением вида $z = f(x, y)$, где $f(x, y)$ — функция, непрерывная в замкнутой квадратируемой области — проекции соответствующего куска поверхности S на плоскость XOY , либо является цилиндрической поверхностью с образующими, параллельными оси OZ , то интеграл по выбранной (любой) стороне поверхности S от непрерывной на поверхности функции $R(x, y, z)$:

$$\iint_S R(x, y, z) \, dx \, dy$$

существует и равен алгебраической сумме двойных интегралов по проекциям указанных кусков поверхности S вида $z = f(x, y)$ на плоскость XOY ¹⁾.

Справедливость этого утверждения вытекает из проведенного при выводе формулы (1) доказательства и свойств поверхностных интегралов. Аналогичные утверждения имеют место для интегралов вида

$$\iint_S P(x, y, z) \, dy \, dz \text{ и } \iint_S Q(x, y, z) \, dz \, dx.$$

Пример. Вычислить интеграл $\iint_S x^2 \, dy \, dz + z^2 \, dx \, dy$ по нижней стороне части конуса $x^2 + y^2 = z^2$, $0 \leq z \leq 1$ (рис. 426).

По определению

$$\iint_S x^2 \, dy \, dz + z^2 \, dx \, dy = \iint_S x^2 \, dy \, dz + \iint_S z^2 \, dx \, dy.$$

Так как поверхность S задается уравнениями $x = \pm \sqrt{z^2 - y^2}$, то для вычисления интеграла $\iint_S x^2 \, dy \, dz$ разбиваем ее плоскостью YOZ на две части: $S_1 \{x = +\sqrt{z^2 - y^2}\}$ и $S_2 \{x = -\sqrt{z^2 - y^2}\}$. При этом

$$\iint_S x^2 \, dy \, dz = \iint_{S_1} x^2 \, dy \, dz + \iint_{S_2} x^2 \, dy \, dz.$$

Проектируя S_1 и S_2 на плоскость YOZ в треугольник σ (рис. 426), по формуле (2) получаем:

$$\iint_{S_1} x^2 \, dy \, dz + \iint_{S_2} x^2 \, dy \, dz = \iint_{\sigma} (\sqrt{z^2 - y^2})^2 \, dy \, dz - \iint_{\sigma} (-\sqrt{z^2 - y^2})^2 \, dy \, dz = 0.$$

¹⁾ По каждому цилиндрическому куску S с образующими, параллельными оси OZ , интеграл $\iint_S R(x, y, z) \, dx \, dy$ равен нулю.

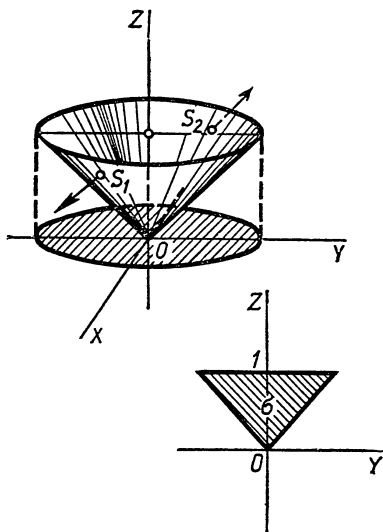


Рис. 426.

Примечание. Знак минус перед вторым интегралом появляется в результате перемены стороны поверхности S_2 .

Вычислим теперь интеграл $\iint_S z^2 dx dy$ по нижней стороне поверхности S . Так как поверхность S можно задать уравнением $z = \sqrt{x^2 + y^2}$, то, меняя сторону поверхности и применяя формулу (1), получаем:

$$\iint_S z^2 dx dy = - \iint_{\sigma'} (x^2 + y^2) dx dy = - \int_0^{2\pi} d\varphi \int_0^1 \rho^3 d\rho = -\frac{\pi}{2}$$

(здесь σ' — проекция S на плоскость XOY).

$$\iint_S x^2 dy dz + z^2 dx dy = -\frac{\pi}{2}.$$

В случае, если поверхность S может быть задана уравнением $z = f(x, y)$, где функция $f(x, y)$ и ее частные производные $f'_x(x, y)$ и $f'_y(x, y)$ непрерывны в замкнутой квадратуемой области σ — проекции поверхности S на плоскость XOY , а функции $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$ непрерывны на поверхности S , имеет место следующая формула:

$$\begin{aligned} \iint_S P dy dz + Q dz dx + R dx dy = \\ = \iint_{\sigma} (-pP - qQ + R) dx dy, \end{aligned} \quad (4)$$

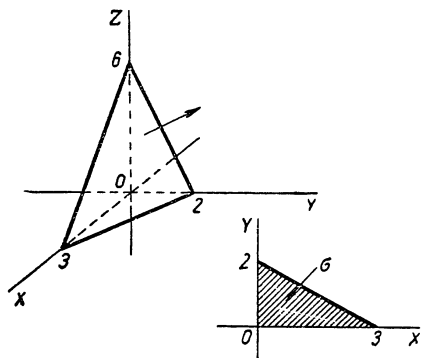


Рис. 427.

где $p = \frac{\partial z}{\partial x}$, $q = \frac{\partial z}{\partial y}$, выражающая поверхностный интеграл общего вида по верхней стороне поверхности S через двойной интеграл по проекции поверхности S на плоскость XOY .

Доказательство формулы (4) приведено в конце следующего параграфа.

Пример. Вычислить интеграл $\iint_S -x dy dz + z dz dx + 5 dx dy$ по верхней стороне лежащей в первом октанте части плоскости $2x + 3y + z = 6$ (рис. 427).

Так как для поверхности S

$$z = 6 - 2x - 3y, \quad p = -2, \quad q = -3,$$

то по формуле (4) получаем:

$$\begin{aligned} \iint_S -x dy dz + z dz dx + 5 dx dy = \iint_{\sigma} [2x - 3(6 - 2x - 3y) + 5] dx dy = \\ = \iint_{\sigma} (8x + 9y - 13) dx dy, \end{aligned}$$

где σ — проекция S на плоскость XOY

Вычисляя двойной интеграл, находим:

$$\begin{aligned} \iint_S -x dy dz + z dz dx + 5 dx dy = \int_0^3 dx \int_0^{\frac{6-2x}{3}} (8x + 9y - 13) dy = \\ = \int_0^3 \left(-\frac{10}{3}x^2 + \frac{38}{3}x - 8 \right) dx = 3. \end{aligned}$$

Аналогично применяются формулы, получаемые из формулы (4) переменных ролей переменных x , y и z .

§ 25. Связь между поверхностными интегралами первого и второго рода

Пусть S — поверхность, заданная уравнением $z = f(x, y)$, где функции $f(x, y)$, $f'_x(x, y)$, $f'_y(x, y)$ непрерывны в замкнутой квадрируемой области σ — проекции поверхности S на плоскость XOY , а функция $R(x, y, z)$ — непрерывная на поверхности S . Выберем верхнюю сторону поверхности S .

Разобьем поверхность S сетью кусочно-гладких кривых на n частей $\Delta S_1, \Delta S_2, \dots, \Delta S_n$. Проекциями этих линий область σ разобьется на части, обозначенные соответственно $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$.

При этом

$$\begin{aligned}\Delta S_i &= \iint_{\Delta\sigma_i} \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)} \, dx \, dy = \\ &= \sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)} \, \Delta\sigma_i,\end{aligned}$$

где $P_i(\xi_i, \eta_i)$ — некоторая точка области $\Delta\sigma_i$.

Обозначим через $M_i(\xi_i, \eta_i, \zeta_i)$, $\zeta_i = f(\xi_i, \eta_i)$, точку поверхности ΔS_i , соответствующую точке P_i , а через γ_i — острый угол, образованный нормалью к поверхности S в точке M_i с осью OZ . Тогда

$$\Delta S_i = \frac{1}{\cos \gamma_i} \Delta\sigma_i, \quad \Delta\sigma_i = \cos \gamma_i \Delta S_i.$$

Составляя интегральную сумму для поверхностного интеграла второго рода по верхней стороне поверхности S , получим:

$$\begin{aligned}\sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i) (\Delta S_i)_{xy} &= \sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i) \Delta\sigma_i = \\ &= \sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i) \cos \gamma_i \Delta S_i.\end{aligned}$$

Сумма $\sum_{i=1}^n R(\xi_i, \eta_i, \zeta_i) \cos \gamma_i \Delta S_i$ является интегральной суммой для интеграла первого рода по поверхности S от функции $R(x, y, z) \cos \gamma$ [γ — угол, составляемый с осью OZ нормалью в текущей точке $M(x, y, z)$ к поверхности S в выбранную сторону поверхности].

При стремлении шага разбиения к нулю в пределе будем иметь

$$\iint_S R(x, y, z) \, dx \, dy = \iint_S R(x, y, z) \cos \gamma \, dS. \quad (1)$$

Если рассматривается нижняя сторона поверхности S , то $(\Delta S_i)_{xy} = -\Delta\sigma_i = \cos \gamma_i \Delta S_i$, где через γ_i обозначен тупой угол, составляемый нормалью в точке M_i к ΔS_i с осью OZ .

Производя аналогичные рассуждения, получаем снова формулу (1):

$$\iint_S R(x, y, z) \, dx \, dy = \iint_S R(x, y, z) \cos \gamma \, dS$$

(γ — угол, составляемый с осью OZ нормалью к поверхности S в выбранную сторону поверхности).

Если поверхность S — цилиндрическая с образующими, параллельными оси OZ , то формула (1) также имеет место, так как

$$\iint_S R(x, y, z) \, dx \, dy = 0 \quad \text{и} \quad \cos \gamma = 0$$

($\gamma = \pm \frac{\pi}{2}$ — угол, составляемый нормалью к поверхности S с осью OZ).

Если поверхность S состоит из конечного числа кусков указанного

вида, то, складывая левые и правые части равенств (1), составленных для ее кусков, получаем формулу

$$\iint_S R(x, y, z) dx dy = \iint_S R(x, y, z) \cos \gamma dS,$$

где γ — угол, составляемый с осью OZ нормалью в текущей точке (x, y, z) к поверхности S в выбранную сторону поверхности.

Аналогичные формулы при соответствующих условиях имеют место для двух других типов поверхностных интегралов. Суммируя их, получаем формулу для поверхностного интеграла общего вида:

$$\iint_S P dy dz + Q dz dx + R dx dy = \iint_S (P \cos \alpha + Q \cos \beta + R \cos \gamma) dS. \quad (2)$$

Здесь $\cos \alpha, \cos \beta, \cos \gamma$ — направляющие косинусы нормали к поверхности S в выбранную (для интеграла слева) сторону поверхности.

Используя связь между поверхностными интегралами первого и второго рода, легко установить справедливость приведенной в конце предыдущего параграфа формулы (4):

$$\iint_S P dy dz + Q dz dx + R dx dy = \iint_{\sigma} (-pP - qQ + R) dx dy,$$

выражающей поверхностный интеграл второго рода по верхней стороне поверхности $S \{z = f(x, y)\}$ через двойной интеграл по проекции этой поверхности на плоскость XOY .

Действительно, так как рассматривается верхняя сторона поверхности S , то направляющие косинусы нормали к поверхности в выбранную сторону соответственно равны:

$$\cos \alpha = \frac{-p}{\sqrt{1+p^2+q^2}}, \quad \cos \beta = \frac{-q}{\sqrt{1+p^2+q^2}}, \quad \cos \gamma = \frac{+1}{\sqrt{1+p^2+q^2}}$$

(см. гл. VII, § 7). Здесь $p = f'_x(x, y)$, $q = f'_y(x, y)$.

При этом по формуле (2)

$$\iint_S P dy dz + Q dz dx + R dx dy = \iint_S \frac{-pP - qQ + R}{\sqrt{1+p^2+q^2}} dS.$$

Выражая поверхностный интеграл первого рода через двойной интеграл по области σ — проекции поверхности S на плоскость XOY (см. (1), § 20), получим требуемое равенство:

$$\iint_S P dy dz + Q dz dx + R dx dy = \iint_{\sigma} (-pP - qQ + R) dx dy.$$

§ 25. Формула Гаусса — Остроградского

Формула Гаусса — Остроградского является аналогом формулы Грина — Остроградского. В то время как формула Грина — Остроградского связывает криволинейный интеграл второго рода по замкнутой кривой с двойным интегралом по плоской области, ограниченной этой кривой,

формула Гаусса — Остроградского устанавливает связь между поверхностным интегралом (второго рода) по замкнутой поверхности и тройным интегралом по пространственной области, ограниченной этой поверхностью.

Рассмотрим пространственную область V , ограниченную снизу и сверху поверхностями S_1 и S_2 , заданными уравнениями $z = h(x, y)$, $z = H(x, y)$ [$h(x, y)$ и $H(x, y)$ — непрерывные функции, $h(x, y) \leq H(x, y)$], сбоку — цилиндрической поверхностью S_3 , образующие которой параллельны оси OZ .

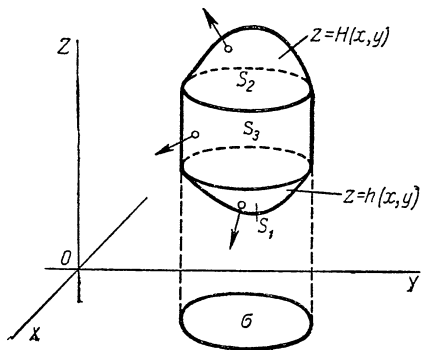


Рис. 428.

Предположим, что эта область V на плоскость XOY проектируется в квадрируемую область σ (рис. 428). Пусть функция $R(x, y, z)$ и ее частная производная $\frac{\partial R(x, y, z)}{\partial z}$ непрерывны в замкнутой области V .

Тогда существует тройной интеграл $\iiint_V \frac{\partial R(x, y, z)}{\partial z} dx dy dz$, причем

$$\begin{aligned} \iiint_V \frac{\partial R(x, y, z)}{\partial z} dx dy dz &= \iint_{\sigma} dx dy \int_{h(x, y)}^{H(x, y)} \frac{\partial R(x, y, z)}{\partial z} dz = \\ &= \iint_{\sigma} R(x, y, H(x, y)) dx dy - \iint_{\sigma} R(x, y, h(x, y)) dx dy. \end{aligned}$$

Двойной интеграл $\iint_{\sigma} R(x, y, H(x, y)) dx dy$ может быть заменен равным ему поверхностным интегралом $\iint_{S_2} R(x, y, z) dx dy$ по верхней стороне поверхности S_2 [см. (1), § 24].

Аналогично интеграл $\iint_{\sigma} R(x, y, h(x, y)) dx dy$ может быть заменен равным ему интегралом $\iint_{S_1} R(x, y, z) dx dy$ по верхней стороне поверхности S_1 .

Следовательно,

$$\iiint_V \frac{\partial R(x, y, z)}{\partial z} dx dy dz = \iint_{S_2} R(x, y, z) dx dy - \iint_{S_1} R(x, y, z) dx dy.$$

Меняя в интеграле по S_1 сторону поверхности и добавляя равный нулю интеграл $\iint_{S_3} R(x, y, z) dx dy$ по «внешней» стороне поверхности S_3 , получим:

$$\begin{aligned} \iiint_V \frac{\partial R(x, y, z)}{\partial z} dx dy dz &= \iint_{S_2} R dx dy + \\ &+ \iint_{S_1} R dx dy + \iint_{S_3} R dx dy = \iint_S R dx dy, \end{aligned}$$

где S — поверхность, ограничивающая область V . Интегрирование ведется по внешней стороне поверхности S .

Таким образом,

$$\iiint_V \frac{\partial R(x, y, z)}{\partial z} dx dy dz = \iint_S R(x, y, z) dx dy. \quad (1)$$

Формула (1) выведена для областей специального вида. Легко видеть, что формула остается справедливой для любой области V , которую можно разбить на конечное число областей рассмотренного выше вида. Применяя к каждой из областей деления формулу (1) и складывая результаты, получим слева тройной интеграл по всей области V , а справа — поверхностный интеграл по внешней стороне поверхности S , ограничивающей область V , так как интегралы по дополнительным «перегородкам» берутся по противоположным сторонам и взаимно уничтожаются.

Аналогичным образом (при соответствующих условиях) выводятся формулы:

$$\iiint_V \frac{\partial P(x, y, z)}{\partial x} dx dy dz = \iint_S P(x, y, z) dy dz, \quad (2)$$

$$\iiint_V \frac{\partial Q(x, y, z)}{\partial y} dx dy dz = \iint_S Q(x, y, z) dz dx. \quad (3)$$

Складывая формулы (1), (2), (3)¹⁾, приходим к так называемой *формуле Гаусса — Остроградского*:

$$\iiint_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx dy dz = \iint_S P dy dz + Q dz dx + R dx dy.$$

Она выражает поверхностный интеграл общего вида по внешней стороне замкнутой поверхности S через тройной интеграл по трехмерной области V , ограниченной этой поверхностью (рис. 429).

Рассмотренные выше формулы (1), (2) и (3) являются ее частными случаями.

Формулу Гаусса — Остроградского можно использовать для вычисления поверхностных интегралов по замкнутым поверхностям.

П р и м е р. Вычислить интеграл $\iint_S x^3 dy dz + y^3 dz dx + z^3 dx dy$ по внешней стороне сферы $x^2 + y^2 + z^2 = R^2$.

Применяя формулу Гаусса — Остроградского, получаем:

$$\begin{aligned} & \iint_S x^3 dy dz + y^3 dz dx + z^3 dx dy = \\ &= \iiint_V (3x^2 + 3y^2 + 3z^2) dx dy dz = \\ &= 3 \int_0^{2\pi} d\varphi \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \theta d\theta \int_0^R \rho^4 d\rho = \frac{12\pi R^5}{5}. \end{aligned}$$

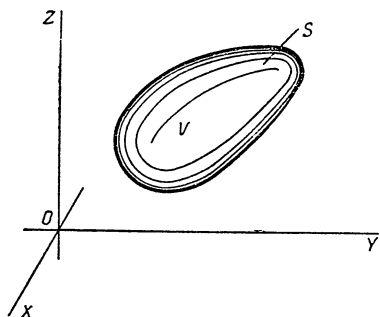


Рис. 429.

В качестве одного из применений формулы Гаусса — Остроградского рассмотрим задачу о вычислении объема тела с помощью поверхностного интеграла по внешней стороне поверхности, ограничивающей это тело.

Действительно, если область V имеет указанную выше форму, то, полагая

$$P(x, y, z) = x, \quad Q(x, y, z) = 0, \quad R(x, y, z) = 0,$$

по формуле Гаусса — Остроградского находим:

$$\iint_S x dy dz = \iiint_V dx dy dz = V.$$

Меняя ролями x, y, z , получим также, что

$$\iint_S y dz dx = V, \quad \iint_S z dx dy = V.$$

Таким образом, имеют место формулы:

$$V = \iint_S x dy dz, \quad V = \iint_S y dz dx, \quad V = \iint_S z dx dy,$$

выражающие объем тела через интеграл по внешней стороне его поверхности.

Взяв функции $P(x, y, z) = \frac{1}{3}x$, $Q(x, y, z) = \frac{1}{3}y$, $R(x, y, z) = \frac{1}{3}z$, получим формулу, выражающую объем тела через поверхностный интеграл общего вида:

$$V = \frac{1}{3} \iint_S x dy dz + y dz dx + z dx dy.$$

¹⁾ Предполагается, что область V удовлетворяет одновременно всем условиям, при которых эти формулы справедливы, а функции $P, \frac{\partial P}{\partial x}, Q, \frac{\partial Q}{\partial y}, R, \frac{\partial R}{\partial z}$ непрерывны в замкнутой области V .

§ 27. Формула Стокса

Формула Стокса устанавливает связь между криволинейным и поверхностным интегралами (второго рода).

Пусть S — поверхность, заданная уравнением $z = f(x, y)$, где функции $f(x, y)$, $f'_x(x, y)$, $f'_y(x, y)$ непрерывны в замкнутой квадратуемой области σ — проекции S на плоскость XOY , L — кусочно-гладкий контур поверхности S , а l — его проекция на плоскость XOY (l — контур области σ) (рис. 430). Выберем на поверхности S определенную сторону, например верхнюю, и установим соответственно положительное направление обхода контура L (см. § 22).

Пусть $x = x(t)$, $y = y(t)$, $\alpha \leq t \leq \beta$, — параметрические уравнения контура l области σ , причем изменению параметра t от α к β соответствует положительное направление обхода контура l ¹⁾.

Тогда уравнения $x = x(t)$, $y = y(t)$, $z = f(x(t), y(t))$, т. е. $z = z(t)$, $\alpha \leq t \leq \beta$, будут параметрическими уравнениями контура L , причем изменению параметра t от α к β будет соответствовать также положительное направление обхода контура L .

Предположим, что функции $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ непрерывны на поверхности S вместе со своими частными производными первого порядка.

Применяя формулу, выражающую криволинейный интеграл через определенный, получим:

$$\begin{aligned} & \oint_L P(x, y, z) dx + Q(\dots) dy + R(\dots) dz = \\ & = \int_{\alpha}^{\beta} [P(x(t), y(t), z(t)) x'(t) + Q(\dots) y'(t) + R(\dots) z'(t)] dt = \\ & = \int_{\alpha}^{\beta} \{ P(x(t), y(t), f(x(t), y(t))) x'(t) + Q(\dots) y'(t) + \\ & \quad + R(\dots) \left[\frac{\partial z}{\partial x} x'(t) + \frac{\partial z}{\partial y} y'(t) \right] \} dt = \\ & = \int_{\alpha}^{\beta} \{ [P(x(t), y(t), f(x(t), y(t))) + pR(\dots)] x'(t) + \\ & \quad + [Q(\dots) + qR(\dots)] y'(t) \} dt = \oint_l [P(x, y, f(x, y)) + \\ & \quad + pR(\dots)] dx + [Q(x, y, f(x, y)) + qR(\dots)] dy, \end{aligned}$$

где

$$p = \frac{\partial z}{\partial x} = f'_x(x, y), \quad q = \frac{\partial z}{\partial y} = f'_y(x, y).$$

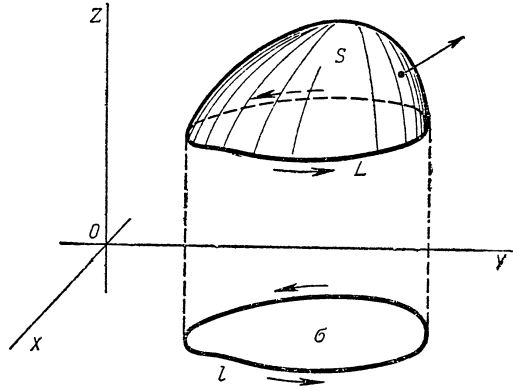


Рис. 430.

¹⁾ Так как l — кусочно-гладкий контур, функции $x(t)$, $y(t)$ можно считать удовлетворяющими условиям, сформулированным в сноске 1) на стр. 509.

По формуле Грина — Остроградского этот криволинейный интеграл может быть заменен равным ему двойным интегралом:

$$\oint_C \{ [Q(x, y, f(x, y)) + qR(\dots)]_x - [P(x, y, f(x, y)) + pR(\dots)]_y \} dx dy.$$

Так как

$$[Q(x, y, f(x, y)) + qR(\dots)]'_x = \frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z} p + \frac{\partial q}{\partial x} R + q \frac{\partial R}{\partial x} + q \frac{\partial R}{\partial z} p,$$

$$[P(x, y, f(x, y)) + pR(\dots)]'_y = \frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} q + \frac{\partial p}{\partial y} R + p \frac{\partial R}{\partial y} + p \frac{\partial R}{\partial z} q,$$

находим:

$$\oint_L P dx + Q dy + R dz = \iint_S \left[-p \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) - q \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \right] dx dy.$$

Интеграл в правой части равенства равен поверхностному интегралу общего вида по верхней стороне поверхности S [см. формулу (4) § 24]:

$$\iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dz dx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

Таким образом, устанавливается следующее соотношение между криволинейным и поверхностным интегралами:

$$\begin{aligned} \oint_L P dx + Q dy + R dz = \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dz dx + \\ + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy^1. \end{aligned} \quad (1)$$

Если переменить сторону поверхности S и соответственно изменить направление обхода контура L , то оба интеграла в соотношении (1) изменят знак. Отсюда следует, что соотношение (1) справедливо и в том случае, когда поверхностный интеграл берется по нижней стороне поверхности S , а криволинейный — по контуру L в положительном на выбранной стороне поверхности направлении.

Применяя свойства криволинейных и поверхностных интегралов, заключаем, что формула

$$\begin{aligned} \oint_L P dx + Q dy + R dz = \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy dz + \\ + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dz dx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy \end{aligned}$$

имеет место для любой незамкнутой, ограниченной контуром L поверхности S , состоящей из конечного числа поверхностей рассмотренного вида, а также для поверхностей, обладающих указанными свойствами,

¹⁾ При доказательстве этого соотношения для простоты предполагалось, что $\frac{\partial p}{\partial y}$, т. е. $\frac{\partial^2 z}{\partial x \partial y}$, и $\frac{\partial q}{\partial x}$, т. е. $\frac{\partial^2 z}{\partial y \partial x}$, существуют и равны: $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$.

относительно других координатных плоскостей. Эта формула называется *формулой Стокса*. Она выражает криволинейный интеграл по контуру L через интеграл по поверхности S , «натянутой» на этот контур. Сторона поверхности и направление обхода контура L взаимно определяют друг друга (см. § 22).

Частным случаем формулы Стокса (L — кривая в плоскости XOY , S — область плоскости XOY , ограниченная этой кривой) является формула Грина — Остроградского:

$$\oint_L P dx + Q dy = \iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

Формулу Стокса можно использовать для вычисления криволинейных интегралов по замкнутым кусочно-гладким контурам.

Пример. Вычислим по формуле Стокса интеграл

$$\oint_L xy dx + y^2 dy + z dz,$$

если L — окружность, заданная уравнениями $x^2 + y^2 = 1$, $z = 1$, а в качестве поверхности S взят параболоид $2 - z = x^2 + y^2$, натянутый на этот контур (рис. 431).

Выбрав на параболоиде верхнюю сторону поверхности, по формуле Стокса получим:

$$\oint_L xy dx + y^2 dy + z dz = \iint_S -x dx dy = - \iint_D x dx dy = 0.$$

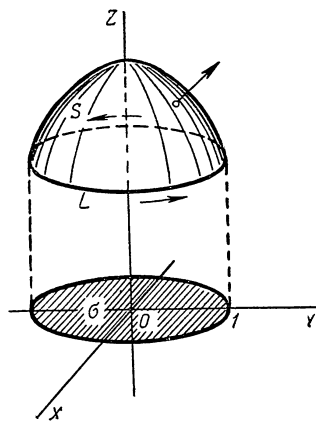


Рис. 431.

ГЛАВА X

РЯДЫ

Раздел 1

ЧИСЛОВЫЕ РЯДЫ

§ 1. Общие сведения

Если члены бесконечной числовой последовательности $u_1, u_2, \dots, u_n, \dots$ соединить знаком $+$, то получится выражение

$$u_1 + u_2 + \dots + u_n + \dots \quad (1)$$

или сокращенно $\sum_{n=1}^{\infty} u_n$, называемое *числовым рядом*. Числа $u_1, u_2, \dots, u_n, \dots$ называются *членами ряда*, u_n — *общим членом ряда*.

Назовем конечные суммы $S_n = u_1 + u_2 + \dots + u_n = \sum_{k=1}^n u_k$ ($n = 1, 2, 3, \dots$) *частными* или *частичными суммами ряда*.

Рассмотрим последовательность $S_1, S_2, \dots, S_n, \dots$ частных сумм ряда (1).

Ряд (1) называется *сходящимся*, если сходится последовательность его частных сумм; *сумма сходящегося ряда* — предел последовательности его частных сумм.

Таким образом, по определению, если $\lim S_n = S$, то S — сумма ряда и в этом случае записывают: $S = \sum_{n=1}^{\infty} u_n$.

Если же последовательность частных сумм ряда расходится, т. е. она или не имеет предела или ее предел равен бесконечности, то ряд называется *расходящимся* и ему не приписывают никакой суммы.

Пример 1. Ряд $a + aq + aq^2 + \dots + aq^{n-1} + \dots$ называется рядом геометрической прогрессии со знаменателем q . Будем рассматривать случай, когда $a \neq 0$.

В этом случае

$$S_n = a + aq + \dots + aq^{n-1} = a \frac{1 - q^n}{1 - q} = \frac{a}{1 - q} - \frac{aq^n}{1 - q},$$

тогда:

а) при $|q| < 1$ $\lim S_n = \frac{a}{1 - q}$, т. е. ряд сходится и его сумма равна $\frac{a}{1 - q}$:

$$\frac{a}{1 - q} = a + aq + aq^2 + \dots + aq^{n-1} + \dots,$$

откуда, например, при $a = 1, q = \frac{1}{2}$, имеем:

$$2 = 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}} + \dots;$$

- в) при $|q| > 1$ ряд расходится, так как $\lim S_n = \infty$;
 с) при $q = 1$ и $q = -1$, получаем соответственно ряды:

$$a + a + a + \dots + a + \dots$$

$$a - a + a - \dots + (-1)^{n+1}a + \dots$$

Первый ряд расходится, так как $S_n = an \rightarrow \infty$. Второй ряд также расходится, так как

$$S_n = \begin{cases} a & \text{при } n \text{ нечетном,} \\ 0 & \text{при } n \text{ четном,} \end{cases}$$

и, следовательно, последовательность $\{S_n\}$ не имеет предела.

Окончательно получаем, что ряд геометрической прогрессии сходится только при $|q| < 1$ и его сумма равна $\frac{a}{1-q}$.

Пример 2. Ряд $1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots$ называется *гармоническим*. Как известно, возрастающая последовательность $\{x_n\}$, $x_n = \left(1 + \frac{1}{n}\right)^n$ сходится и ее предел равен e :

$$\lim \left(1 + \frac{1}{n}\right)^n = e,$$

при этом $e > \left(1 + \frac{1}{n}\right)^n$. Логарифмируя это неравенство, запишем:

$$1 > n \ln \frac{n+1}{n},$$

откуда

$$\frac{1}{n} > \ln(n+1) - \ln n.$$

Полагая $n = 1, 2, \dots$, получаем:

$$1 > \ln 2$$

$$\frac{1}{2} > \ln 3 - \ln 2$$

$$\frac{1}{3} > \ln 4 - \ln 3$$

$$\dots\dots\dots$$

$$\frac{1}{n} > \ln(n+1) - \ln n.$$

Сложив эти неравенства, получим:

$$S_n > \ln(n+1).$$

Следовательно, $S_n \rightarrow \infty$, гармонический ряд расходится.

Заметим, что любую последовательность $u_1, u_2, \dots, u_n, \dots$ можно рассматривать как последовательность частных сумм ряда

$$u_1 + (u_2 - u_1) + (u_3 - u_2) + \dots + (u_n - u_{n-1}) + \dots$$

Если последовательность $\{u_n\}$ сходится, то ее предел и есть сумма этого ряда.

Свойства рядов.

1. Сходимость или расходимость ряда не изменится (поведение ряда относительно сходимости не изменится), если изменить, отбросить или добавить конечное число членов ряда.

Это следует из того, что две последовательности, члены которых, начиная с некоторого, отличаются между собой на одно и то же число, ведут себя одинаково относительно сходимости, т. е. или обе сходятся, или обе расходятся.

2. Рассмотрим ряд $\sum u_n$ (для простоты записи пределы суммирования опускаем) и ряд, полученный умножением каждого члена ряда $\sum u_n$ на число c , т. е. ряд $\sum cu_n$.

Теорема. Если ряд $\sum u_n$ сходится и его сумма равна S , то и ряд $\sum cu_n$ также сходится и его сумма равна cS .

Действительно, пусть S_n — частные суммы ряда $\sum u_n$, тогда $S_n \rightarrow S$ и, следовательно, $cS_n \rightarrow cS$, но cS_n — частные суммы ряда $\sum cu_n$, поэтому $\sum cu_n = cS$.

Заметим: если ряд $\sum u_n$ расходится, то и ряд $\sum cu_n$, где $c \neq 0$, также расходится. В самом деле, если допустить сходимость ряда $\sum cu_n$ и умножить его члены на $\frac{1}{c}$, то получится сходящийся ряд $\sum u_n$, что противоречит условию.

3. Рассмотрим два ряда $\sum u_n$ и $\sum v_n$. Ряды $\sum (u_n \pm v_n)$, полученные соответственно сложением и вычитанием членов этих рядов, имеющих одинаковые номера, или, что то же, полученные почленным сложением и вычитанием рядов $\sum u_n$ и $\sum v_n$, называют соответственно *суммой* и *разностью данных рядов*.

Теорема. Если ряды $\sum u_n$ и $\sum v_n$ сходятся и их суммы соответственно равны S и σ , то и ряд $\sum (u_n + v_n)$ сходится и его сумма равна $S + \sigma$:

$$\sum (u_n + v_n) = \sum u_n + \sum v_n.$$

В самом деле, пусть S_n и σ_n — частные суммы соответствующих рядов, тогда по условию $S_n \rightarrow S$, $\sigma_n \rightarrow \sigma$ и, следовательно, $S_n + \sigma_n \rightarrow S + \sigma$. Учитывая, что $S_n + \sigma_n$ — частные суммы ряда $\sum (u_n + v_n)$, получим:

$$\sum (u_n + v_n) = S + \sigma.$$

Таким образом, сумма двух сходящихся рядов является сходящимся рядом.

Разность двух сходящихся рядов $\sum u_n$ и $\sum v_n$ есть ряд сходящийся, так как ряд $\sum (u_n - v_n)$ является суммой двух сходящихся рядов $\sum u_n$ и $\sum -v_n$.

Сумма сходящегося и расходящегося ряда — расходящийся ряд. Действительно, пусть ряд $\sum u_n$ сходится, а ряд $\sum v_n$ расходится. Тогда, допустив сходимость ряда $\sum (u_n + v_n)$, получим, что ряд $\sum v_n$ сходится [как разность двух сходящихся рядов $\sum (u_n + v_n)$ и $\sum u_n$], а это противоречит условию.

Общего утверждения относительно суммы двух расходящихся рядов сделать нельзя. В одних случаях сумма двух расходящихся рядов может быть сходящимся рядом, в других — расходящимся рядом.

4. При рассмотрении числовых рядов практически решаются две задачи: 1) исследовать, сходится или расходится ряд, другими словами, исследовать ряд на сходимость; 2) зная, что ряд сходится, найти его сумму.

Обе эти задачи не являются простыми. Как первая, так и вторая задачи практически не всегда выполнимы.

Учитывая, что любая частная сумма сходящегося ряда дает приближенное значение его суммы, на практике часто заменяют сумму ряда его частной суммой S_n . При этом производят оценку погрешности, происходящей от такой замены, точнее, оценивают разность $R_n = S - S_n$, называемую *остатком ряда*.

Мы будем решать главным образом первую задачу, т. е. будем исследовать ряд на сходимость.

Рассмотрим различные признаки сходимости рядов.

Критерий Коши. Сначала рассмотрим критерий Коши для последовательности, т. е. признак Коши, *необходимый и достаточный* для сходимости последовательности.

Для того чтобы числовая последовательность $a_1, a_2, \dots, a_n, \dots$ была сходящейся, *необходимо и достаточно*, чтобы для любого числа $\varepsilon > 0$ существовал такой номер N , что при $n > N$ и любом целом $p > 0$ выполнялось бы неравенство $|a_{n+p} - a_n| < \varepsilon$.

Доказательство необходимости. Пусть $a_n \rightarrow a$, тогда для любого числа $\varepsilon > 0$ найдется такой номер N , что неравенство $|a - a_n| < \frac{\varepsilon}{2}$ выполняется при $n > N$. При $n > N$ и любом целом $p > 0$ выполняется также неравенство

$$|a - a_{n+p}| < \frac{\varepsilon}{2}.$$

Учитывая оба неравенства, получаем:

$$\begin{aligned} |a_{n+p} - a_n| &= |(a_{n+p} - a) + (a - a_n)| \leq \\ &\leq |a_{n+p} - a| + |a - a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

следовательно, $|a_{n+p} - a_n| < \varepsilon$ при $n > N$ и любом целом $p > 0$, что и требовалось доказать.

Доказательство достаточности. По условию для любого числа $\varepsilon > 0$ существует натуральное число N такое, что при $n > N$ и любом целом $p > 0$ выполняется неравенство

$$|a_{n+p} - a_n| < \varepsilon$$

или, что то же, выполняются неравенства

$$a_n - \varepsilon < a_{n+p} < a_n + \varepsilon. \quad (2)$$

Возьмем $n = N + 1$ и будем p давать последовательно значения, равные 1, 2, ... Тогда из неравенства (2) получим, что все члены последовательности, кроме, может быть, конечного числа их, принадлежат отрезку $[a_n - \varepsilon, a_n + \varepsilon]$ длины 2ε .

Таким образом, для каждого числа $\varepsilon > 0$ найдется отрезок длины 2ε , которому принадлежат все члены последовательности, кроме, может быть, конечного числа их.

Пусть значения ε образуют последовательность $1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$. Тогда для этой последовательности значений ε определится соответствующая ей последовательность отрезков $\Delta_1, \Delta_2, \dots, \Delta_n, \dots$, длины которых стремятся к нулю (так как длина Δ_n равна $\frac{2}{n}$).

Рассмотрим теперь последовательность отрезков $\Delta'_1, \Delta'_2, \dots, \Delta'_n, \dots$ образованную по следующему закону: $\Delta'_1 = \Delta_1$, Δ'_2 — общая часть Δ_1 и Δ_2 , $\dots$, Δ'_n — общая часть $\Delta_1, \Delta_2, \dots, \Delta_n$ и т. д. (Δ'_n может вырождаться в точку).

Последовательность $\{\Delta'_n\}$ является последовательностью вложенных стягивающихся отрезков, обладающих тем свойством, что каждый из них содержит все члены последовательности, кроме, может быть, конечного числа их.

По лемме о вложенных стягивающихся отрезках существует точка, обозначим ее a , принадлежащая всем отрезкам. Нетрудно видеть, что число a и есть предел последовательности $\{a_n\}$.

В самом деле, какую бы окрестность точки a ни взяли, в нее попадет, при некотором n , отрезок Δ'_n , а следовательно, попадут все члены последовательности, кроме, может быть, конечного числа их.

Таким образом, $\lim a_n = a$, что и требовалось доказать.

Используя связь между числовыми последовательностями и числовыми рядами, можно сформулировать критерий Коши для ряда:

Для того чтобы ряд $u_1 + u_2 + \dots + u_n + \dots$ был сходящимся, необходимо и достаточно, чтобы для любого числа $\varepsilon > 0$ существовал такой номер N , что при $n > N$ и любом целом $p > 0$ выполнялось бы неравенство

$$|u_{n+1} + u_{n+2} + \dots + u_{n+p}| < \varepsilon$$

или, что то же, чтобы все достаточно далеко лежащие отрезки¹⁾ этого ряда были по модулю как угодно малы.

Действительно, если применить доказанный критерий Коши для

¹⁾ Отрезком ряда называется сумма нескольких подряд взятых членов этого ряда.

последовательности к последовательности частных сумм ряда, а также учесть, что $S_{n+p} - S_n = u_{n+1} + u_{n+2} + \dots + u_{n+p}$, то получим критерий Коши для ряда.

Следует иметь в виду, что пользоваться критерием Коши на практике обычно затруднительно, поэтому рассматривают более простые в смысле их практического приложения, но зато более «узкие» признаки сходимости рядов. Начнем их изучение с необходимых признаков сходимости рядов.

Необходимые признаки сходимости рядов

1. Если ряд $\sum u_n$ сходится, то его общий член стремится к нулю:

$$u_n \rightarrow 0.$$

Действительно, так как ряд $\sum u_n$ сходится, то положив в критерии Коши $p=1$, получим: для любого $\varepsilon > 0$ найдется $N = N(\varepsilon)$, такое, что $|u_{n+1}| < \varepsilon$ при $n > N(\varepsilon)$, а это и означает, что $u_n \rightarrow 0$. Другое доказательство: по условию ряд $\sum u_n$ сходится, обозначим через S его сумму. Тогда $S_n \rightarrow S$ и, следовательно, $u_n = S_n - S_{n-1} \rightarrow S - S = 0$, $u_n \rightarrow 0$.

Заметим, что условие $u_n \rightarrow 0$ не является достаточным для сходимости ряда $\sum u_n$, как это легко видеть на примере гармонического ряда.

Действительно, гармонический ряд $\sum_{n=1}^{\infty} \frac{1}{n}$ расходится, хотя общий член ряда $u_n = \frac{1}{n} \rightarrow 0$.

Таким образом, если общий член ряда $\sum u_n$ не стремится к нулю, то ряд расходится. Если же общий член ряда стремится к нулю, то о сходимости ряда еще ничего сказать нельзя. В этом случае ряд нужно еще исследовать, применив другие признаки сходимости. Некоторые из них будут рассмотрены ниже.

2. Если ряд сходится, то последовательность его частных сумм ограничена.

Действительно, по определению сходящегося ряда последовательность его частных сумм имеет конечный предел, а значит, эта последовательность ограничена.

Заметим, что ограниченности последовательности частных сумм ряда еще недостаточно для сходимости ряда, так как ограниченности последовательности недостаточно для ее сходимости.

Например, ряд $1 - 1 + 1 - 1 + \dots + (-1)^{n+1} + \dots$ расходится, так как последовательность его частных сумм $\{S_n\}$ расходится в силу того, что

$$S_n = \begin{cases} 0 & \text{при } n \text{ четных,} \\ 1 & \text{при } n \text{ нечетных.} \end{cases}$$

Но последовательность частных сумм этого ряда ограничена, $|S_n| < 2$ при любом n .

§ 2. Ряды с неотрицательными членами

При изучении знакопостоянных рядов достаточно рассмотреть ряды с положительными, точнее, с неотрицательными членами (в целях общности мы допускаем и существование членов, равных нулю). Действительно, ряды с отрицательными, точнее, с неположительными членами отличаются от соответствующих рядов с неотрицательными членами только множителем -1 , и вследствие этого такие ряды ведут себя одинаково относительно сходимости. Рассмотрим признаки сходимости *рядов с неотрицательными членами*.

Теорема. Для сходимости ряда $\sum u_n$ с неотрицательными членами необходимо и достаточно, чтобы частные суммы ряда были ограничены.

В самом деле, ограниченность частных сумм ряда является необхо-

димым условием сходимости любого ряда, и в частности, и ряда с неотрицательными членами.

Достаточность условия. Так как ряд $\sum u_n$ имеет неотрицательные члены, то его частные суммы образуют неубывающую последовательность, $S_n \leq S_{n+1}$ при любом n .

По условию последовательность $\{S_n\}$ частных сумм ряда $\sum u_n$ ограничена, а значит, по теореме о монотонных ограниченных последовательностях, она сходится, т. е. сходится ряд $\sum u_n$.

Например, ряд

$$\frac{1}{2+1^2} + \frac{1}{2^2+2^2} + \frac{1}{2^3+3^2} + \dots + \frac{1}{2^n+n^2} + \dots$$

сходится, так как $\frac{1}{2^n+n^2} < \frac{1}{2^n}$ ($n=1, 2, \dots$) и $S_n < \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^n} < 1$, $S_n < 1$ при всех n .

Заметим, что ряд с членами произвольных знаков может расходиться по двум причинам: или частные суммы ряда не имеют предела, как, например, для ряда $\sum (-1)^{n+1}$, или предел частных сумм ряда равен бесконечности, как, например, для ряда $\sum q^n$, $|q| > 1$.

Для ряда с неотрицательными членами в силу монотонности последовательности его частных сумм первое утверждение невозможно, так как последовательность частных сумм такого ряда всегда имеет предел, конечный или бесконечный. Таким образом, если ряд с неотрицательными членами расходится, то последовательность его частных сумм имеет бесконечный предел.

Признак (принцип) сравнения рядов с неотрицательными членами. Пусть даны два ряда с неотрицательными членами:

$$u_1 + u_2 + u_3 + \dots + u_n + \dots, \quad u_n \geq 0, \quad (1)$$

$$v_1 + v_2 + v_3 + \dots + v_n + \dots, \quad v_n \geq 0. \quad (2)$$

Теорема. Если $u_n \leq v_n$ при любом n , то из сходимости ряда (2) следует сходимость ряда (1), а из расходимости ряда (1) следует расходимость ряда (2).

Доказательство. Обозначим через S_n и σ_n соответственно частные суммы рядов (1) и (2). Так как ряд (1) сходится, то его частные суммы ограничены, т. е. при всех n $\sigma_n < M$, где M — некоторое число. Но $S_n \leq \sigma_n$, поэтому частные суммы ряда (2) также ограничены, а этого, как известно, достаточно для сходимости ряда (2) с неотрицательными членами.

Если же ряд (1) расходится, то ряд (2) также расходится, так как, допустив сходимость ряда (2) получим по только что доказанному сходимость ряда (1), а это противоречит условию.

Таким образом, теорема полностью доказана.

Пример 1. Ряд $\frac{1}{\ln 2} + \frac{1}{\ln 3} + \dots + \frac{1}{\ln n} + \dots$ расходится, так как $\frac{1}{\ln n} > \frac{1}{n}$, а гармонический ряд $\sum_{n=1}^{\infty} \frac{1}{n}$ расходится.

Пример 2. Ряд $\sum_{n=1}^{\infty} \frac{1}{n2^n}$ сходится, так как сходится ряд геометрической прогрессии $\sum_{n=1}^{\infty} \frac{1}{2^n}$, а члены данного ряда не больше соответствующих членов этого ряда сходящейся геометрической прогрессии:

$$\frac{1}{n2^n} \leq \frac{1}{2^n}.$$

Если $u_n > 0$, $v_n > 0$ и существует

где h — число, отличное от нуля, то ряды $\sum u_n$ и $\sum v_n$ ведут себя одинаково относительно сходимости.

В самом деле, так как число h — отличное от нуля ($h > 0$), то существуют положительные числа h_1 и h_2 , удовлетворяющие условиям $h_1 < h < h_2$. Тогда по определению предела последовательности для всех n , кроме возможно конечного числа их, выполняются неравенства:

$$h_1 < \frac{u_n}{v_n} < h_2$$

или неравенства

$$h_1 v_n < u_n < h_2 v_n.$$

Если ряд $\sum v_n$ сходится, то сходится и ряд $\sum h_2 v_n$ и, следовательно, если учесть правое неравенство, сходится ряд $\sum u_n$.

Если же ряд $\sum v_n$ расходится, то расходится и ряд $\sum h_1 v_n$ и, если учесть левое неравенство, расходится и ряд $\sum u_n$.

Например, ряд $\frac{1}{a+h} + \frac{1}{a+2h} + \dots + \frac{1}{a+nh} + \dots$ ($h > 0$) ($a > 0$) рас-
ходится, так как

$$\frac{1}{n} : \frac{1}{a + nh} = \frac{a + nh}{n} \rightarrow h, \quad h > 0.$$

Признак сравнения рядов может быть использован не только для исследования конкретно задаваемых рядов, а также и для получения практически удобных общих признаков сходимости рядов.

Признак Даламбера ¹⁾. Если для ряда $\sum u_n$ с положительными членами существует такое число $q < 1$, что для всех достаточно больших n выполняется неравенство $\frac{u_{n+1}}{u_n} \leq q$, то ряд $\sum u_n$ сходится, если же для всех достаточно больших n $\frac{u_{n+1}}{u_n} \geq 1$, то ряд $\sum u_n$ расходится.

Доказательство. Пусть, например, для $n \geq t$ выполняется неравенство $\frac{u_{n+1}}{u_n} \leq q$ или, что то же, неравенство $u_{n+1} \leq u_n q$. Тогда

$$\begin{aligned} u_{m+1} &\leq u_m q, \\ u_{m+2} &\leq u_{m+1} q \leq u_m q^2, \\ u_{m+3} &\leq u_{m+2} q \leq u_m q^3, \\ &\vdots \\ &\vdots \end{aligned}$$

Сходимость ряда $\sum u_n$ вытекает из сходимости ряда геометрической прогрессии $u_m q + u_m q^2 + \dots + u_m q^k + \dots$ со знаменателем q , $0 < q < 1$.

Если же $\frac{u_{n+1}}{u_n} \geq 1$, то при достаточно больших n $u_{n+1} \geq u_n$, т. е. каждый последующий член данного ряда, начиная с некоторого номера, не меньше предыдущего, поэтому общий член ряда u_n при $n \rightarrow \infty$ не стремится к нулю, а значит, ряд $\sum u_n$ расходится.

¹⁾ Даламбер (1717—1783) — французский математик и механик.

Следует иметь в виду, что условие $\frac{u_{n+1}}{u_n} \leq q$ ($q < 1$) нельзя заменить условием $\frac{u_{n+1}}{u_n} < 1$. В последнем случае ряд может быть и расходящимся, как это имеет место, например, в случае ряда $\sum_{n=1}^{\infty} \frac{1}{n}$.

Для практического применения удобнее пользоваться следствием из признака Даламбера, или, как его иногда называют, *предельным признаком Даламбера*.

Следствие. Если существует $\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \rho$, то при $\rho < 1$ ряд сходится, при $\rho > 1$ ряд расходится, при $\rho = 1$ признак вопроса не решает. (Заметим, что случай $\rho = +\infty$ включается в $\rho > 1$.)

Доказательство. Если $\rho < 1$, то возьмем число q , удовлетворяющее условию $\rho < q < 1$. Тогда для всех значений n , кроме, возможно, конечного числа, будет выполняться неравенство

$$\frac{u_{n+1}}{u_n} < q \quad (q < 1),$$

и, следовательно, по признаку Даламбера ряд $\sum u_n$ сходится. Если же $\rho > 1$, то для всех n , начиная с некоторого номера, выполняется неравенство $\frac{u_{n+1}}{u_n} > 1$, следовательно, по признаку Даламбера ряд $\sum u_n$ расходится.

Пример 1. Ряд $\sum_{n=1}^{\infty} \frac{n}{2^n}$ сходится, так как

$$u_n = \frac{n}{2^n} \text{ и } \lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \frac{(n+1)2^n}{2^{n+1}n} = \frac{1}{2} < 1.$$

Пример 2. Ряд $1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!} + \dots$ сходится, так как в этом случае $u_n = \frac{1}{n!}$ и

$$\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!} = \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0 < 1.$$

Пример 3. Для каждого из рядов $\sum_{n=1}^{\infty} \frac{1}{n}$, $\sum_{n=1}^{\infty} \frac{1}{n^2}$, как это нетрудно проверить, $\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = 1$. При этом ряд $\sum_{n=1}^{\infty} \frac{1}{n}$ расходится, а ряд $\sum_{n=1}^{\infty} \frac{1}{n^2}$, как это будет показано в § 3, сходится.

Признак Коши. Если существует такое число $q < 1$, что для всех достаточно больших n выполняется неравенство $\sqrt[n]{u_n} \leq q$, то ряд $\sum u_n$ сходится, если же для всех достаточно больших n $\sqrt[n]{u_n} \geq 1$, то ряд $\sum u_n$ расходится.

Доказательство. Условие $u_n \leq q^n$ при $0 < q < 1$ означает, что все члены ряда $\sum u_n$, кроме конечного числа их, не превосходят членов ряда сходящейся геометрической прогрессии $\sum q^n$, а следовательно, ряд $\sum u_n$ сходится.

Если $\sqrt[n]{u_n} \geq 1$, то $u_n \geq 1$, т. е. не выполнено необходимое условие сходимости ряда, ряд расходится.

Следствие. Если существует $\lim_{n \rightarrow \infty} \sqrt[n]{u_n} = \rho$, то при $\rho < 1$ ряд сходится, при $\rho > 1$ ряд расходится, при $\rho = 1$ признак вопроса не решает (случай $\rho = +\infty$ включается в $\rho > 1$).

Доказательство. Пусть $\rho < 1$. Возьмем число q , удовлетворяющее условию $\rho < q < 1$. Тогда для всех n , кроме конечного числа, выполняется неравенство $u_n < q^n$ и, следовательно, по признаку Коши, ряд $\sum u_n$ сходится.

Если же $\rho > 1$, то начиная с некоторого номера имеем $\sqrt[n]{u_n} > 1$ или $u_n > 1$, следовательно, не выполняется необходимое условие сходимости ряда, ряд $\sum u_n$ расходится.

Пример. Ряд $\sum_{n=1}^{\infty} \left(\frac{2n^2+1}{3n^2+5} \right)^n$ сходится, так как

$$\lim_{n \rightarrow \infty} \sqrt[n]{u_n} = \lim_{n \rightarrow \infty} \frac{2n^2+1}{3n^2+5} = \frac{2}{3} < 1.$$

Пример 2. Рассмотрим ряд $\sum_{n=1}^{\infty} \left(1 + \frac{1}{n} \right)^n$. Для него

$$\lim_{n \rightarrow \infty} \sqrt[n]{u_n} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right) = 1,$$

поэтому вопрос о сходимости данного ряда признаком Коши не решается. Нетрудно видеть, что этот ряд расходится, так как

$$u_n = \left(1 + \frac{1}{n} \right)^n \rightarrow e \neq 0.$$

Пример 3. Ряд $\sum_{n=1}^{\infty} \frac{1}{n^2}$ сходится, как это будет показано в § 3, хотя, так же как и в предшествующем примере,

$$\lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n^2}} = \lim_{n \rightarrow \infty} \left[\left(\frac{1}{n} \right)^{\frac{1}{n}} \right]^2 = 1$$

(так как $\lim_{x \rightarrow +0} x^x = 1$, см. § 14, гл. IV).

Как видно из предшествующих рассуждений, признаки Даламбера и Коши основаны на сравнении данного ряда с рядом геометрической прогрессии. Эти признаки не являются «чувствительными» к рядам, сходящимся «медленнее», чем геометрическая прогрессия. Для таких рядов рассматривают более сильные признаки, в частности интегральный признак, с которым мы познакомимся в следующем параграфе.

§ 3. Несобственные интегралы

В § 15, гл. VI были рассмотрены несобственные интегралы с бесконечными пределами

$$\int_a^{+\infty} f(x) dx, \quad \int_{-\infty}^b f(x) dx, \quad \int_{-\infty}^{+\infty} f(x) dx,$$

где $f(x)$ — непрерывная функция соответственно на промежутках $[a, +\infty)$, $(-\infty, b]$, $(-\infty, +\infty)$, а также интегралы вида $\int_a^b f(x) dx$, где $f(x)$ непрерывная, неограниченная функция на промежутках $(a, b]$, $[a, b)$, (a, b) .

Будем изучать сходимость интегралов с бесконечными пределами вида $\int_a^{+\infty} f(x) dx$. Аналогичные рассуждения можно провести и для интегралов $\int_{-\infty}^b f(x) dx$ и $\int_{-\infty}^{+\infty} f(x) dx$.

Как известно, несобственный интеграл $\int_a^{+\infty} f(x) dx$ называется *сходящимся*, если интеграл $\int_a^b f(x) dx$, $a \leq b < +\infty$, при $b \rightarrow +\infty$ имеет конечный предел. Если же интеграл $\int_a^b f(x) dx$ при $b \rightarrow +\infty$ не имеет предела или имеет бесконечный предел, то несобственный интеграл $\int_a^{+\infty} f(x) dx$ называется *расходящимся*. Если интеграл $\int_a^{+\infty} f(x) dx$ сходится, то его значение принимается равным пределу интеграла $\int_a^b f(x) dx$ при $b \rightarrow +\infty$:

$$\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} \int_a^b f(x) dx.$$

Во многих задачах, связанных с несобственными интегралами, достаточно только выяснить вопрос о сходимости рассматриваемого интеграла и не требуется находить его значение.

Исследование на сходимость интеграла вида $\int_a^{+\infty} f(x) dx$ имеет много общего с исследованием ряда на сходимость.

Действительно, по аналогии с рядами интеграл $\int_a^b f(x) dx$, $a \leq b < +\infty$, называется *частным*, или *частичным*, *интегралом несобственного интеграла* $\int_a^{+\infty} f(x) dx$. И тогда определение сходимости интеграла $\int_a^{+\infty} f(x) dx$ принимает следующий вид: интеграл $\int_a^{+\infty} f(x) dx$ сходится, если сходятся его частные интегралы, т. е. если частные интегралы имеют конечный предел при $b \rightarrow +\infty$. Если же частные интегралы расходятся, т. е. не имеют предела или их предел бесконечен, то интеграл $\int_a^{+\infty} f(x) dx$ называется *расходящимся*.

Некоторые свойства несобственных интегралов.

1) Интегралы $\int_a^{+\infty} f(x) dx$ и $\int_{a_1}^{+\infty} f(x) dx$, $a_1 > a$, при условии, что $f(x)$ — непрерывная функция на промежутке $[a, +\infty)$, ведут себя одинаково относительно сходимости. Это следует из того, что частные интегралы $\int_a^b f(x) dx$ и $\int_{a_1}^b f(x) dx$ при любом b , $b > a_1$, отличаются между собой на одно и то же число.

2) Если интеграл $\int_a^{+\infty} f(x) dx$ сходится и равен I , то интеграл $\int_a^{+\infty} cf(x) dx$, где c — число, также сходится и равен cI . Или, что то же, постоянный множитель можно вносить под знак сходящегося несобственного интеграла:

$$c \int_a^{+\infty} f(x) dx = \int_a^{+\infty} cf(x) dx.$$

В самом деле, по условию $\int_a^b f(x) dx \rightarrow I$, следовательно, $\int_a^b cf(x) dx = c \int_a^b f(x) dx \rightarrow cI$, что и требовалось доказать.

Заметим, что если интеграл $\int_a^{+\infty} f(x) dx$ расходится, то и интеграл $\int_a^{+\infty} cf(x) dx$ (где $c \neq 0$) также расходится.

Действительно, допустив сходимость интеграла $\int_a^{+\infty} cf(x) dx$, получим при умножении подинтегральной функции на $\frac{1}{c}$ сходимость интеграла $\int_a^{+\infty} f(x) dx$, а это противоречит условию.

3) Если интегралы $\int_a^{+\infty} f(x) dx$ и $\int_a^{+\infty} f_1(x) dx$ сходятся и соответственно равны I и I_1 , то интеграл $\int_a^{+\infty} [f(x) + f_1(x)] dx$ также сходится и равен $I + I_1$. Или, что то же, несобственный интеграл от суммы двух функций равен сумме несобственных интегралов от слагаемых, если сходится интеграл от каждого слагаемого:

$$\int_a^{+\infty} [f(x) + f_1(x)] dx = \int_a^{+\infty} f(x) dx + \int_a^{+\infty} f_1(x) dx.$$

Действительно, по условию

$$\begin{aligned} \int_a^b f(x) dx &\rightarrow I, \\ \int_a^b f_1(x) dx &\rightarrow I_1, \end{aligned}$$

откуда, следовательно,

$$\int_a^b [f(x) + f_1(x)] dx = \int_a^b f(x) dx + \int_a^b f_1(x) dx \rightarrow I + I_1.$$

Заметим, что если интеграл $\int_a^{+\infty} f(x) dx$ сходится, а интеграл $\int_a^{+\infty} f_1(x) dx$ расходится, то интеграл $\int_a^{+\infty} [f(x) + f_1(x)] dx$ расходится, так как в противном случае сходил бы интеграл $\int_a^{+\infty} f_1(x) dx$, будучи разностью двух сходящихся интегралов.

Если же интегралы $\int_a^{+\infty} f(x) dx$ и $\int_a^{+\infty} f_1(x) dx$ расходятся, то о сходимости интеграла $\int_a^{+\infty} [f(x) + f_1(x)] dx$ в общем случае ничего сказать нельзя, в одних случаях он может быть сходящимся, в других — расходящимся.

Критерий Коши. Для сходимости интеграла $\int_a^{+\infty} f(x) dx$ необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало такое число $A = A(\varepsilon)$, что при любых p и q , $p > A$, $q > A$, выполнялось бы неравенство

$$\left| \int_p^q f(x) dx \right| < \varepsilon.$$

Доказательство необходимости. В самом деле, так как $\int_a^b f(x) dx \rightarrow I$, то для любого числа $\varepsilon > 0$ найдется $A = A(\varepsilon)$, такое, что неравенство $\left| I - \int_a^b f(x) dx \right| < \frac{\varepsilon}{2}$ выполняется при любом $b > A$. В частности, это неравенство выполняется и при $b = p$ и $b = q$, где $p > A$, $q > A$. Тогда имеем:

$$\begin{aligned} \left| \int_p^q f(x) dx \right| &= \left| \int_a^q f(x) dx - \int_a^p f(x) dx \right| = \left| \left[\int_a^q f(x) dx - I \right] + \left[I - \int_a^p f(x) dx \right] \right| \leq \\ &\leq \left| \int_a^q f(x) dx - I \right| + \left| I - \int_a^p f(x) dx \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Окончательно получаем, что неравенство $\left| \int_p^q f(x) dx \right| < \varepsilon$ выполняется при любых $p > A$, $q > A$.

Доказательство достаточности мы опускаем.

Необходимое условие сходимости несобственных интегралов. Если интеграл $\int_a^{+\infty} f(x) dx$ сходится, то его частные интегралы ограничены, т. е. существует такое число $M > 0$, что для всех $b > a$ выполняется неравенство

$$\left| \int_a^b f(x) dx \right| < M.$$

Это утверждение непосредственно вытекает из существования конечного предела частных интегралов.

Таким образом, если частные интегралы не ограничены, то несобственный интеграл расходится. Если же частные интегралы ограничены, то о сходимости несобственного интеграла еще ничего сказать нельзя: в одних случаях интеграл может сходиться, а в других —

расходиться. Например, у несобственного интеграла $\int_0^{+\infty} \cos x dx$ частные интегралы $\int_0^b \cos x dx = \sin b$ ограничены, но они не имеют предела при $b \rightarrow +\infty$ (так как $\sin b$ при $b \rightarrow +\infty$ не имеет предела) и, следовательно,

интеграл $\int_0^{+\infty} \cos x \, dx$ — расходящийся. Аналогично получаем, что интеграл $\int_0^{+\infty} \sin x \, dx$ расходится.

Несобственные интегралы от неотрицательных функций.

Теорема. Если подинтегральная функция $f(x)$ на промежутке $[a, +\infty)$ непрерывна и неотрицательна, $f(x) \geq 0$, то ограниченность частных интегралов является *необходимым и достаточным* условием сходимости интеграла

$$\int_a^{+\infty} f(x) \, dx.$$

Необходимость условия была только что доказана для интегралов от любых непрерывных функций.

Если же $f(x) \geq 0$, то это условие будет и достаточным, так как в этом случае при $b \rightarrow +\infty$ частные интегралы $\int_a^b f(x) \, dx$ возрастают и, оставаясь ограниченными, стремятся к конечному пределу¹⁾.

Принцип сравнения несобственных интегралов.

Теорема. Если функции $f(x)$ и $\varphi(x)$ непрерывны на полуинтервале $[a, +\infty)$ и удовлетворяют на нем условию $0 \leq f(x) \leq \varphi(x)$, то из сходимости интеграла

$$\int_a^{+\infty} \varphi(x) \, dx \tag{1}$$

следует сходимость интеграла

$$\int_a^{+\infty} f(x) \, dx, \tag{2}$$

а из расходимости интеграла (2) следует расходимость интеграла (1).

Доказательство этой теоремы аналогично доказательству принципа сравнения числовых рядов.

В самом деле, так как интеграл (1) сходится, то его частные интегралы ограничены. Но так как

$$\int_a^b f(x) \, dx \leq \int_a^b \varphi(x) \, dx,$$

то, следовательно, ограничены и частные интегралы несобственного интеграла (2). А этого, как известно, достаточно для сходимости интеграла (2) от неотрицательной функции. Если же интеграл (2) расходится, то, допустив сходимость интеграла (1), получим по только что доказанному сходимость интеграла (2), что противоречит условию. Окончательно получаем, что интеграл (1) также расходится.

Следствие. Если $f(x)$ и $\varphi(x)$ непрерывные функции на полуинтервале $[a, +\infty)$ и

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{\varphi(x)} = h \tag{3}$$

(h — число, $h \neq 0$), то интегралы $\int_a^{+\infty} f(x) \, dx$ и $\int_a^{+\infty} \varphi(x) \, dx$ ведут себя одинаково относительно сходимости.

¹⁾ Теорема существования предела монотонной ограниченной последовательности может быть обобщена на случай монотонной ограниченной функции.

Действительно, из условия (3) следует, что функции $f(x)$ и $\varphi(x)$ отличны от нуля при достаточно больших x . Если учесть первые два свойства несобственных интегралов, то станет ясно, что, не ограничивая общности рассуждений, можно предположить, что число $h > 0$, а функции $f(x)$ и $\varphi(x)$ положительны при $x \geq a$.

Это следствие доказывается аналогично тому, как это делалось в теории рядов (см. § 2). Действительно, если выбрать h_1 и h_2 так, чтобы выполнялись условия $0 < h_1 < h < h_2$, и если учесть, что для всех достаточно больших значений x справедливы неравенства

$$h_1 < \frac{f(x)}{\varphi(x)} < h_2,$$

а значит, неравенства

$$h_1 \varphi(x) < f(x) < h_2 \varphi(x),$$

то на основании доказанной теоремы легко получить требуемое утверждение.

Предлагается самостоятельно провести эти рассуждения.

Пример 1. Интеграл вида $\int_a^{+\infty} \frac{dx}{x^\alpha}$ ($\alpha > 0$) сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$. (Этот пример уже рассматривался в гл. VI, но в силу его особой значимости повторим еще раз приведенные выше рассуждения.)

Действительно, при $\alpha \neq 1$

$$\int_a^b \frac{dx}{x^\alpha} = \frac{x^{-\alpha+1}}{-\alpha+1} \Big|_a^b = \frac{b^{-\alpha+1}}{-\alpha+1} - \frac{a^{-\alpha+1}}{-\alpha+1}$$

Если учесть, что $\lim_{b \rightarrow +\infty} \frac{b^{-\alpha+1}}{-\alpha+1} = \begin{cases} 0 & \text{при } \alpha > 1, \\ +\infty & \text{при } \alpha < 1, \end{cases}$ то получим сходимость интеграла $\int_a^{+\infty} \frac{dx}{x^\alpha}$ при $\alpha > 1$ и расходимость при $\alpha < 1$.

При $\alpha = 1$ имеем $\int_a^b \frac{dx}{x} = \ln x \Big|_a^b = \ln b - \ln a \rightarrow +\infty$ при $b \rightarrow +\infty$ и, следовательно, интеграл $\int_a^{+\infty} \frac{dx}{x}$ расходится.

Пример 2. Рассмотрим интеграл от рациональной дроби $\int_a^{+\infty} R(x) dx$, где $R(x) = \frac{P(x)}{Q(x)}$, $P(x)$ и $Q(x)$ — многочлены, причем $Q(x)$ не имеет корней при $x \geq a$.
Так как

$$R(x) = \frac{a_0 x^n + a_1 x^{n-1} + \dots + a_n}{b_0 x^m + b_1 x^{m-1} + \dots + b_m} = \frac{R_1(x)}{x^{m-n}}, \text{ где } R_1(x) = \frac{a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n}}{b_0 + \frac{b_1}{x} + \dots + \frac{b_m}{x^m}} \rightarrow \frac{a_0}{b_0}$$

при $x \rightarrow +\infty$, то учитывая пример 1, получим, что интеграл $\int_a^{+\infty} R(x) dx$ сходится лишь при условии, что $m - n \geq 2$.

Пример 3. Интеграл $\int_a^{+\infty} \frac{dx}{x (\ln x)^\alpha}$ ($\alpha > 1$) сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$. Этот пример решается так же, как и пример 1. Действительно, при $\alpha \neq 1$

$$\int_a^b \frac{dx}{x (\ln x)^\alpha} = \frac{(\ln x)^{-\alpha+1}}{-\alpha+1} \Big|_a^b = \frac{(\ln b)^{-\alpha+1}}{-\alpha+1} - \frac{(\ln a)^{-\alpha+1}}{-\alpha+1}.$$

Но так как

$$\lim_{b \rightarrow +\infty} \frac{(\ln b)^{-\alpha+1}}{-\alpha+1} = \begin{cases} 0 & \text{при } \alpha > 1, \\ +\infty & \text{при } \alpha < 1, \end{cases}$$

то $\int_a^{+\infty} \frac{dx}{x (\ln x)^\alpha}$ сходится при $\alpha > 1$ и расходится при $\alpha < 1$.

Если же $\alpha = 1$, то $\int_a^b \frac{dx}{x \ln x} = \ln \ln b - \ln \ln a \rightarrow +\infty$ при $b \rightarrow +\infty$, следовательно, $\int_a^{+\infty} \frac{dx}{x \ln x}$ расходится.

Пусть теперь $f(x)$ — произвольная непрерывная на промежутке $[a, +\infty)$ функция, не обязательно знакопостоянная. Интеграл $\int_a^{+\infty} f(x) dx$ называется *абсолютно сходящимся*, если сходится интеграл $\int_a^{+\infty} |f(x)| dx$.

Теорема. Если сходится интеграл $\int_a^{+\infty} |f(x)| dx$, то сходится и интеграл $\int_a^{+\infty} f(x) dx$.

Действительно, применяя критерий Коши к интегралу $\int_a^{+\infty} |f(x)| dx$, получим: для любого $\varepsilon > 0$ существует число $A(\varepsilon)$, такое, что при любых p и q , $p < q$, удовлетворяющих условию $p > A$, $q > A$, справедливо неравенство

$$\left| \int_p^q |f(x)| dx \right| < \varepsilon, \text{ а значит, и неравенство } \left| \int_p^q f(x) dx \right| < \varepsilon.$$

Но $\left| \int_p^q f(x) dx \right| \leq \int_p^q |f(x)| dx$ и, следовательно, $\left| \int_p^q f(x) dx \right| < \varepsilon$ при $p > A$, $q > A$, откуда по критерию Коши получаем сходимость интеграла $\int_a^{+\infty} f(x) dx$, что и доказывает теорему.

Заметим, что сходимость несобственного интеграла от знакопеременной функции не влечет за собой его абсолютной сходимости, а для интегралов от знакопостоянных функций из сходимости интеграла следует его абсолютная сходимость. Например, можно показать, что интеграл $\int_a^{+\infty} \frac{\sin x}{x} dx$ ($a > 0$) сходится, а интеграл $\int_a^{+\infty} \left| \frac{\sin x}{x} \right| dx$ расходится, т. е.

интеграл $\int_a^{+\infty} \frac{\sin x}{x} dx$ не является абсолютно сходящимся. Интеграл $\int_a^{+\infty} \frac{\sin x}{x^2} dx$ ($a > 0$) — сходится абсолютно, так как $\left| \frac{\sin x}{x^2} \right| \leq \frac{1}{x^2}$ и интеграл $\int_a^{+\infty} \frac{dx}{x^2}$ сходится.

Интегральный признак сходимости рядов с положительными членами. Если $\varphi(x)$ непрерывная положительная убывающая функция на промежутке $[1, +\infty)$, то ряд $\varphi(1) + \varphi(2) + \dots + \varphi(n) + \dots = \sum_{n=1}^{\infty} \varphi(n)$ и интеграл $\int_1^{+\infty} \varphi(x) dx$ ведут себя одинаково относительно сходимости.

Доказательство. Учитывая геометрический смысл определенного интеграла (рис. 432), запишем:

$$\varphi(2) + \varphi(3) + \dots + \varphi(n) < \int_1^n \varphi(x) dx < \varphi(1) + \varphi(2) + \dots + \varphi(n-1)$$

или

$$S_n - \varphi(1) < \int_1^n \varphi(x) dx < S_{n-1},$$

где S_n — частные суммы рассматриваемого ряда.

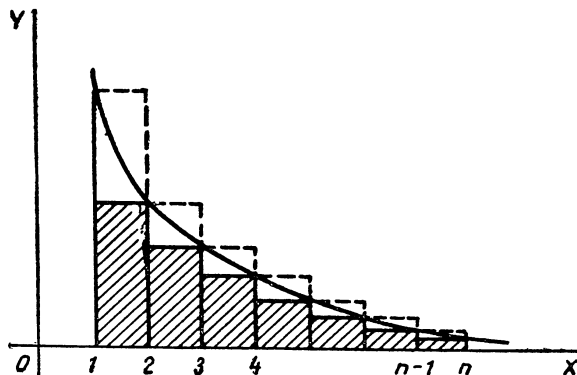


Рис. 432.

Если интеграл $\int_1^{+\infty} \varphi(x) dx$ сходится, то его частные интегралы $\int_1^n \varphi(x) dx$ ограничены. Учитывая левое неравенство, получаем, что частные суммы ряда также ограничены, а этого, как известно, достаточно для сходимости ряда с положительными членами, т. е. ряд $\sum_{n=1}^{\infty} \varphi(n)$ сходится.

Если же интеграл $\int_1^{+\infty} \varphi(x) dx$ расходится, то $\int_1^n \varphi(x) dx \rightarrow +\infty$ при $n \rightarrow \infty$ (как монотонно возрастающая неограниченная последовательность), и, учитывая правое неравенство, получим, что $S_{n-1} \rightarrow +\infty$, т. е. ряд $\sum_{n=1}^{\infty} \varphi(n)$ расходится.

Пример 1. Ряд $1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \dots + \frac{1}{n^\alpha} + \dots$ сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$ (так как $\int_1^{+\infty} \frac{dx}{x^\alpha}$ сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$).

В частности, при $\alpha = 2$ получим сходящийся ряд

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{n^2} + \dots,$$

при $\alpha = 1$ получим известный расходящийся гармонический ряд, при $\alpha = \frac{1}{2}$ получим расходящийся ряд

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} \dots + \frac{1}{\sqrt{n}} + \dots$$

Пример 2. Ряд $\sum_{n=2}^{\infty} \frac{1}{n (\ln n)^{\alpha}}$ сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$. Это следует из интегрального признака сходимости рядов, если учесть, что интеграл $\int_2^{\infty} \frac{dx}{x (\ln x)^{\alpha}}$ сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$.

В частности, при $\alpha = 1$ получаем расходящийся ряд

$$\frac{1}{2 \ln 2} + \frac{1}{3 \ln 3} + \dots + \frac{1}{n \ln n} + \dots$$

И, наконец, сделаем несколько замечаний о сходимости несобственных интегралов от неограниченных функций. Не ограничивая общности рассуждений, рассмотрим несобственные интегралы вида $\int_a^b f(x) dx$, где $f(x)$ — непрерывная функция на полуинтервале $(a, b]$, неограниченная вблизи a . Теория этих интегралов аналогична теории несобственных интегралов с бесконечными пределами. В частности, принцип сравнения для таких интегралов можно сформулировать следующим образом.

Если функции $f(x)$ и $\varphi(x)$ непрерывны на полуинтервале $(a, b]$ и для всех точек x , лежащих на полуинтервале $(a, b]$ вблизи точки a , выполняются условия $0 \leq f(x) \leq \varphi(x)$, то из сходимости интеграла

$$\int_a^b \varphi(x) dx \quad (1)$$

следует сходимость интеграла

$$\int_a^b f(x) dx, \quad (2)$$

а из расходимости интеграла (2) следует расходимость интеграла (1).

Следствие. Если $f(x)$ и $\varphi(x)$ непрерывные функции на полуинтервале $(a, b]$ и $\lim_{x \rightarrow a+0} \frac{f(x)}{\varphi(x)} = h$, где h — число, отличное от нуля, $h \neq 0$, то интегралы $\int_a^b f(x) dx$ и $\int_a^b \varphi(x) dx$ ведут себя одинаково относительно сходимости.

Пример 1. Интеграл $\int_a^b \frac{dx}{(x-a)^{\alpha}}$ сходится при $\alpha < 1$ и расходится при $\alpha \geq 1$.

В самом деле, при $\alpha \neq 1$

$$\int_{a+\varepsilon}^b \frac{dx}{(x-a)^{\alpha}} = \frac{(x-a)^{-\alpha+1}}{-\alpha+1} \Big|_{a+\varepsilon}^b = \frac{(b-a)^{-\alpha+1}}{-\alpha+1} - \frac{\varepsilon^{-\alpha+1}}{-\alpha+1}.$$

Так как

$$\lim_{\varepsilon \rightarrow 0, \varepsilon > 0} \frac{\varepsilon^{-\alpha+1}}{-\alpha+1} = \begin{cases} 0 & \text{при } \alpha < 1, \\ -\infty & \text{при } \alpha > 1, \end{cases}$$

то $\int_a^b \frac{dx}{(x-a)^{\alpha}}$ сходится при $\alpha < 1$ и расходится при $\alpha > 1$.

При $x=1$ получим $\int_{a+\varepsilon}^b \frac{dx}{(x-a)} = \ln(b-a) - \ln \varepsilon \rightarrow +\infty$ при $\varepsilon \rightarrow +0$, т. е. интеграл $\int_a^b \frac{dx}{(x-a)}$ расходится.

Пример 2. Интеграл $\int_0^1 \frac{dx}{x^2(2-x^2)}$ расходится, так как $\frac{1}{x^2(2-x^2)} : \frac{1}{x^2} = \frac{1}{2-x^2} \rightarrow \frac{1}{2}$ при $x \rightarrow 0$.

§ 4. Знакопеременные ряды

Прежде чем рассматривать ряды с членами произвольных знаков, рассмотрим их частный случай, а именно ряды, члены которых имеют чередующиеся знаки, такие ряды называются *знакопередающимися*.

Знакопередающиеся ряды. Знакопередающийся ряд, если первый член положителен, можно записать в виде:

$$u_1 - u_2 + u_3 - u_4 + \dots + (-1)^{n+1} u_n + \dots, \text{ где } u_n > 0, n = 1, 2, 3, \dots$$

Признак Лейбница. Если у знакопередающегося ряда

$$u_1 - u_2 + u_3 - u_4 + \dots + (-1)^{n+1} u_n + \dots \quad (1)$$

абсолютные величины членов ряда убывают: $u_1 > u_2 > u_3 > \dots$ и общий член стремится к нулю, $u_n \rightarrow 0$, то ряд сходится.

Доказательство. Рассмотрим частные суммы ряда с четным числом членов:

$$S_{2n} = u_1 - u_2 + u_3 - u_4 + \dots + u_{2n-1} - u_{2n} = (u_1 - u_2) + (u_3 - u_4) + \dots + (u_{2n-1} - u_{2n}). \quad (2)$$

Так как каждая из скобок, в силу первого условия, положительна, то последовательность $\{S_{2n}\}$ возрастающая. Докажем, что она ограничена. Для этого представим S_{2n} следующим образом:

$$S_{2n} = u_1 - [(u_2 - u_3) + (u_4 - u_5) + \dots + (u_{2n-2} - u_{2n-1}) + u_{2n}].$$

Число, записанное в квадратных скобках, положительно, поэтому

$$S_{2n} < u_1.$$

Итак, последовательность $\{S_{2n}\}$ монотонно возрастающая, ограниченная и, следовательно, сходящаяся.

Пусть $\lim_{n \rightarrow \infty} S_{2n} = S$.

Чтобы доказать сходимость ряда (1), нужно доказать еще, что последовательность частных сумм нечетного числа членов этого ряда также сходится и имеет предел, равный S .

Так как $S_{2n+1} = S_{2n} + u_{2n+1}$ и $u_{2n+1} \rightarrow 0$ (по условию), то $\lim_{n \rightarrow \infty} S_{2n+1} = S$, т. е. последовательность всех частных сумм ряда (1) сходится и, значит, ряд (1) сходится. Таким образом, теорема Лейбница доказана.

Заметим, что для суммы S ряда (1) справедливо соотношение $0 < S < u_1$. Действительно, частные суммы четного номера S_{2n} приближаются к сумме S , возрастаая, следовательно, $S > S_{2n}$ при любом n . Кроме того, $S_{2n} > 0$ ($n = 1, 2, \dots$), как это следует из равенства (2), а значит, и $S > 0$. Частные суммы нечетного номера S_{2n+1} можно записать в виде:

$$S_{2n+1} = u_1 - (u_2 - u_3) - \dots - (u_{2n} - u_{2n+1}).$$

Отсюда видно, что последовательность $\{S_{2n+1}\}$ монотонно убывающая и что $S_{2n+1} < u_1$ при любом n . Так как $S < S_{2n+1}$ при любом n , то, следовательно, $S < u_1$. Суммируя сказанное, получаем: $0 < S < u_1$.

Рассмотрим пример. Ряд $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + (-1)^{n+1} \frac{1}{n} + \dots$ сходится, по признаку Лейбница. Этот ряд отличается от гармонического ряда только знаками членов четных номеров.

Отметим следующее свойство знакочередующегося ряда, (удовлетворяющего условиям признака Лейбница), имеющее большое практическое приложение.

Пусть ряд $u_1 - u_2 + u_3 - u_4 + \dots + (-1)^{n+1} u_n + \dots$ сходится и его сумма равна S , тогда

$$S = S_n \pm (u_{n+1} - u_{n+2} + u_{n+3} - \dots). \quad (2^1)$$

Разность $S - S_n = R_n$ — остаток ряда, в свою очередь, является суммой знакочередующегося ряда и, следовательно, удовлетворяет условию $|R_n| < u_{n+1}$. Таким образом, заменяя сумму ряда его частной суммой, получаем ошибку, абсолютная величина которой меньше абсолютной величины первого отброшенного члена ряда.

Например, ряд $1 - \frac{1}{3 \cdot 3} + \frac{1}{5 \cdot 3^3} - \frac{1}{7 \cdot 3^5} + \dots + (-1)^{n+1} \frac{1}{(2n-1) 3^{n-1}} + \dots$ сходится по признаку Лейбница. Если положить его сумму¹⁾ S , приближенно равной сумме первых шести членов этого ряда, то получим ошибку, абсолютная величина которой меньше, чем

$$\frac{1}{13 \cdot 3^5} < \frac{1}{9477} < 0,001.$$

Абсолютная и условная сходимость рядов. Возьмем какой-нибудь знакпеременный ряд

$$u_1 + u_2 + \dots + u_n + \dots, \quad (3)$$

т. е. ряд с членами произвольных знаков. Рассмотрим ряд, составленных из абсолютных величин членов ряда (3):

$$|u_1| + |u_2| + \dots + |u_n| + \dots \quad (4)$$

Теорема. Если ряд (4) сходится, то сходится и ряд (3).

Доказательство. Пусть ряд (4) сходится. По критерию Коши для любого $\varepsilon > 0$ найдется число N , такое, что при $n > N$ и любом целом $p > 0$ выполняется неравенство

$$|u_{n+1}| + |u_{n+2}| + \dots + |u_{n+p}| < \varepsilon$$

По свойству абсолютных величин:

$$|u_{n+1} + u_{n+2} + \dots + u_{n+p}| \leq |u_{n+1}| + |u_{n+2}| + \dots + |u_{n+p}| < \varepsilon,$$

откуда получаем, что неравенство

$$|u_{n+1} + u_{n+2} + \dots + u_{n+p}| < \varepsilon$$

справедливо при $n > N$ и любом $p > 0$, а значит, по критерию Коши, ряд (1) сходится.

Например, ряд

$$\frac{\sin \frac{\pi}{4}}{3} + \frac{\sin 3 \frac{\pi}{4}}{3^2} + \dots + \frac{\sin (2n-1) \frac{\pi}{4}}{3^n} + \dots$$

не является знакочередующимся, так как за каждыми двумя положительными членами следует два отрицательных члена, значит, признак

¹⁾ Отметим, что $2 \sqrt[3]{3} S = \pi$ (см. пример 2, § 9, при $x = \frac{\sqrt[3]{3}}{3}$).

Лейбница к нему не применим. Но если рассмотреть ряд, составленный из абсолютных величин его членов, то получим ряд геометрической прогрессии

$$\frac{\sqrt{2}}{3 \cdot 2} + \frac{\sqrt{2}}{3^2 \cdot 2} + \dots + \frac{\sqrt{2}}{3^n \cdot 2} + \dots,$$

знаменатель которой $\frac{1}{3} < 1$, он сходится, а значит, по доказанной теореме и данный ряд сходится и притом абсолютно.

Заметим, что из сходимости ряда (3) не следует сходимость ряда (4). Например, ряд $1 - \frac{1}{2} + \frac{1}{3} - \dots + \frac{(-1)^{n+1}}{n} + \dots$ сходится, а гармонический ряд $1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots$ расходится.

Ряд $\sum u_n$ называется *абсолютно сходящимся*, если сходится ряд $\sum |u_n|$, составленный из абсолютных величин его членов. Например, ряд $1 - \frac{1}{2^2} + \frac{1}{3^2} - \dots + \frac{(-1)^{n+1}}{n^2} + \dots$ абсолютно сходящийся, так как ряд $1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{n^2} + \dots$ сходится. Очевидно, для знакопостоянных рядов понятия сходимости и абсолютной сходимости совпадают.

Ряд $\sum u_n$ называется *условно сходящимся*, если он сходится, а ряд $\sum |u_n|$ расходится. Например, ряд $1 - \frac{1}{2^p} + \frac{1}{3^p} - \dots + \frac{(-1)^{n+1}}{n^p} + \dots$ при $p \leq 1$ условно сходящийся, так как он сходится (по признаку Лейбница), а ряд из модулей его членов $1 + \frac{1}{2^p} + \frac{1}{3^p} + \dots + \frac{1}{n^p} + \dots$ расходится.

Следует иметь в виду, что деление сходящихся рядов на абсолютно и условно сходящиеся весьма существенно. Оказывается, что основные свойства конечных сумм переносятся только на абсолютно сходящиеся ряды, тогда как условно сходящиеся ряды некоторыми из этих свойств не обладают.

Признаки Даламбера и Коши для знакопеременных рядов.

Пусть ряд $\sum u_n$ имеет члены произвольных знаков.

Признак Даламбера. Если существует $\lim_{n \rightarrow \infty} \left| \frac{u_{n+1}}{u_n} \right| = \rho$, то при $\rho < 1$ ряд $\sum u_n$ абсолютно сходится, при $\rho > 1$ ряд расходится. При $\rho = 1$ признак вопроса не решает.

В самом деле, если $\rho < 1$, то по признаку Даламбера для рядов с положительными членами ряд $\sum |u_n|$ сходится, а следовательно, ряд $\sum u_n$ абсолютно сходится. Если же $\rho > 1$, то $|u_{n+1}| > |u_n|$ при любом достаточно большом n , поэтому $|u_n|$ не стремится к нулю, а значит, и u_n не стремится к нулю, т. е. ряд $\sum u_n$ расходится.

Если же $\rho = 1$, то в общем случае о сходимости ряда ничего сказать нельзя. Например, $\rho = 1$ для расходящегося гармонического ряда $\sum_{n=1}^{\infty} \frac{1}{n}$, а ряд $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}$ абсолютно сходится, хотя для него также $\rho = 1$.

Признак Коши. Если существует $\lim_{n \rightarrow \infty} \sqrt[n]{|u_n|} = \rho$, то при $\rho < 1$ ряд $\sum u_n$ абсолютно сходится, при $\rho > 1$ ряд расходится. При $\rho = 1$ признак вопроса не решает.

Действительно, если $\rho < 1$, то по признаку Коши для рядов с неотрицательными членами ряд $\sum |u_n|$ сходится следовательно, ряд $\sum u_n$ абсо-

лютно сходится. Если же $\rho > 1$, то при любом достаточно большом n $\sqrt[n]{|u_n|} > 1$, а значит, $|u_n| > 1$. Отсюда получаем, что $|u_n|$, а следовательно, и u_n не стремятся к нулю, т. е. ряд $\sum u_n$ расходится.

Свойства абсолютно сходящихся рядов.

Теорема. Для абсолютной сходимости ряда $\sum u_n$ необходимо и достаточно, чтобы его можно было представить в виде разности двух сходящихся рядов с неотрицательными членами.

Доказательство. Пусть ряд $\sum u_n$ абсолютно сходится. Обозначим через $\sum a_n$ ряд, составленный из его положительных членов и нулей на местах отрицательных членов, а через $\sum b_n$ — ряд, составленный из абсолютных величин отрицательных членов и нулей на местах положительных членов. Тогда $\sum u_n = \sum (a_n - b_n)$. Ряды $\sum a_n$ и $\sum b_n$ с неотрицательными членами по признаку сравнения сходятся, так как выполняются неравенства

$$a_n \leq |u_n|, \quad b_n \leq |u_n| \quad (n = 1, 2, \dots)$$

Таким образом, необходимость условия доказана. Заметим, что в этом случае сумма ряда $\sum u_n$ равна разности сумм рядов

$$\sum a_n \text{ и } \sum b_n, \quad \sum u_n = \sum (a_n - b_n) = \sum a_n - \sum b_n.$$

Теперь докажем, что условие достаточно. Пусть $\sum u_n = \sum (a_n - b_n)$, при этом $\sum a_n$ и $\sum b_n$ — сходящиеся ряды с неотрицательными членами. Тогда $|u_n| = |a_n - b_n| \leq |a_n| + |b_n| = a_n + b_n$, т. е. $|u_n| \leq a_n + b_n$, и если учесть, что ряд $\sum (a_n + b_n)$ сходится (как сумма двух сходящихся рядов), то получим сходимость ряда $\sum |u_n|$, т. е. абсолютную сходимость ряда $\sum u_n$, что и требовалось доказать.

Следствие. Условно сходящийся ряд является разностью двух расходящихся рядов с неотрицательными стремящимися к нулю членами. Действительно, $\sum u_n = \sum (a_n - b_n)$, причем ни один из рядов $\sum a_n$, $\sum b_n$ не является сходящимся (так как разность сходящегося и расходящегося ряда — расходящийся ряд). Из условия сходимости ряда $\sum u_n$ непосредственно следует, что $u_n \rightarrow 0$, а значит, $a_n \rightarrow 0$ и $b_n \rightarrow 0$.

Теорема о группировке членов ряда. В сходящемся ряде любая группировка членов ряда, не изменяющая их порядка, сохраняет сходимость ряда и величину суммы. Действительно, пусть ряд

$$u_1 + u_2 + \dots + u_n + \dots \quad (5)$$

сходится и его сумма равна S . Рассмотрим ряд

$$(u_1 + u_2 + \dots + u_k) + (u_{k+1} + u_{k+2} + \dots + u_l) + (u_{l+1} + u_{l+2} + \dots + u_m) + \dots, \quad (5')$$

полученный из ряда (5) группировкой членов при сохранении их порядка.

Последовательность частных сумм ряда (5') является частью последовательности $\{S_n\}$ частных сумм ряда (5). Но так как ряд (5) сходится, то $S_n \rightarrow S$ и, следовательно, последовательность частных сумм ряда (5') сходится и имеет тот же предел S , т. е. ряд (5') имеет сумму, равную S . Заметим, что из сходимости ряда (5') не всегда следует сходимость ряда (5), т. е. в общем случае опускать скобки нельзя. В этом можно убедиться на примере ряда

$$(1 - 1) + (1 - 1) + \dots + (1 - 1) + \dots$$

Этот ряд сходится, так как его можно записать в следующем виде:

$$0 + 0 + 0 + \dots + 0 + \dots,$$

в то время как ряд

$$1 - 1 + 1 - 1 + \dots + (-1)^{n+1} + \dots$$

расходится.

Не доказывая, заметим, что если члены ряда (5) неотрицательны, $u_n \geq 0$ ($n = 1, 2, \dots$), то из сходимости ряда (5') следует сходимость ряда (5).

Теорема о перестановке членов ряда. Если ряд сходится абсолютно, то ряд, полученный из него любой перестановкой членов, также абсолютно сходится и имеет ту же сумму.

Сначала докажем эту теорему для рядов с неотрицательными членами.

Пусть ряд

$$u_1 + u_2 + u_3 + \dots + u_n + \dots, \quad u_n \geq 0, \quad (6)$$

сходится и его сумма равна S . Произвольным образом переставляя его члены, получим ряд

$$u'_1 + u'_2 + u'_3 + \dots + u'_n + \dots \quad (6')$$

Обозначим через S_n и S'_n соответственно частные суммы рядов (6) и (6').

Рассмотрим S'_n , n — любое число. Так как каждый член ряда (6') является членом ряда (6), то можно взять k настолько большим, чтобы каждый из членов, входящих в S'_n , содержался в S_k .

Нетрудно видеть, что тогда $S'_n \leq S_k \leq S$, т. е. частные суммы ряда (6') с неотрицательными членами ограничены и не превосходят S . Следовательно, ряд (6') сходится и его сумма, которую обозначим S' , не больше S , $S' \leq S$.

Но ряд (6) можно рассматривать как ряд, полученный из ряда (6') в результате перестановки его членов, и, значит, по доказанному $S \leq S'$.

Сравнивая эти неравенства, окончательно получаем $S = S'$, что и требовалось доказать.

Доказательство теоремы для абсолютно сходящихся рядов получается непосредственно. Действительно, пусть ряд $\sum u_n$ сходится и его сумма равна S , а ряд $\sum u'_n$ получен перестановкой членов ряда $\sum u_n$. По доказанной выше теореме ряды $\sum u_n$ и $\sum u'_n$ являются разностями сходящихся рядов с неотрицательными членами: $\sum u_n = \sum (a_n - b_n)$, $\sum u'_n = \sum (a'_n - b'_n)$. $\sum a'_n$ и $\sum b'_n$ — ряды, полученные перестановкой соответственно членов сходящихся рядов $\sum a_n$ и $\sum b_n$ с неотрицательными членами.

Как уже доказано, ряды $\sum a'_n$ и $\sum b'_n$ сходятся и их суммы равны суммам соответствующих рядов $\sum a_n$ и $\sum b_n$, т. е.

$$\sum a'_n = \sum a_n, \quad \sum b'_n = \sum b_n.$$

Следовательно, ряд $\sum u'_n = \sum (a'_n - b'_n)$ абсолютно сходится (как разность двух сходящихся рядов с неотрицательными членами) и его сумма равна сумме ряда $\sum u_n$, так как

$$\sum u'_n = \sum (a'_n - b'_n) = \sum a'_n - \sum b'_n = \sum a_n - \sum b_n.$$

Таким образом, теорема полностью доказана. Заметим, что эта теорема не будет иметь места в случае условно сходящихся рядов. Убедимся в этом на примере. Как известно, ряд

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{(-1)^{n+1}}{n} + \dots \quad (6)$$

сходится условно. Обозначим его сумму через S . В этом ряде сделаем следующую перестановку членов: за каждым положительным членом поставим следующие два отрицательных члена. Получим ряд

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \dots \quad (7')$$

для которого можно доказать, что он сходится. Тогда группируя его члены, получим ряд

$$\left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} + \dots, \quad (8)$$

который также сходится и его сумма равна сумме ряда (7').

Но ряд (8) можно записать в следующем виде:

$$\frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \dots = \frac{1}{2} \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots\right),$$

откуда следует, что его сумма равна $\frac{S}{2}$.

Итак, перестановкой членов ряда (7) мы получили новый ряд, сумма которого в два раза меньше суммы данного ряда.

Как доказал Риман¹⁾, перестановкой членов условно сходящегося ряда можно получить сходящийся ряд, имеющий любую наперед заданную сумму, и даже расходящийся ряд. Доказательства этой теоремы мы не приводим.

И, наконец, формулируем теорему, доказательство которой мы опускаем, но на которую нам придется неоднократно ссылаться.

Теорема. При любой группировке членов абсолютно сходящегося ряда (при этом число групп может быть как конечным так и бесконечным и число членов в каждой группе может быть как конечным так и бесконечным) получается сходящийся ряд, сумма которого равна сумме данного ряда.

Теорема умножения абсолютно сходящихся рядов. Если ряды $\sum_{n=1}^{\infty} u_n$

и $\sum_{n=1}^{\infty} v_n$ сходятся абсолютно и их суммы соответственно равны S и σ ,

то ряд, составленный из всех произведений вида $u_i v_k$ ($i, k=1, 2, \dots$), взятых в каком угодно порядке, также сходится абсолютно и его сумма равна $S \cdot \sigma$ — произведению сумм перемножаемых рядов.

Доказательство этой теоремы мы опускаем.

Ряд, составленный из всевозможных произведений вида

$$u_i v_k \quad (i, k=1, 2, \dots),$$

называют произведением абсолютно сходящихся рядов $\sum u_n$ и $\sum v_n$.

На практике часто умножение рядов производят по следующей схеме:

$$\begin{aligned} \sum_{n=1}^{\infty} u_n \cdot \sum_{n=1}^{\infty} v_n &= (u_1 + u_2 + u_3 + \dots + u_n + \dots) (v_1 + v_2 + v_3 + \dots + \\ &+ \dots + v_n + \dots) = u_1 v_1 + (u_1 v_2 + u_2 v_1) + (u_1 v_3 + u_2 v_2 + u_3 v_1) + \\ &+ \dots + \dots + \dots + (u_1 v_n + u_2 v_{n-1} + \dots + u_n v_1) + \dots \end{aligned}$$

Следует иметь в виду, что если производить умножение условно сходящихся рядов только что указанным способом, то в результате умножения можно получить расходящийся ряд.

¹⁾ Риман (1826—1866) — немецкий математик.

§ 5. Ряды с комплексными членами

Пусть $z_1, z_2, z_3, \dots, z_n, \dots$ ($z_n = x_n + iy_n$) — бесконечная последовательность комплексных чисел. Число $z = x + iy$ называется *пределом последовательности* $\{z_n\}$, если для любого числа $\varepsilon > 0$ найдется такой номер $N = N(\varepsilon)$, что при $n > N$ выполняется неравенство $|z - z_n| < \varepsilon$.

Геометрически это означает, что для любого числа $\varepsilon > 0$, начиная с некоторого номера, зависящего от ε , все члены последовательности попадут в круг радиуса ε с центром в точке z , вне круга останется лишь конечное число членов последовательности. Тот факт, что число z является пределом последовательности, записывают так: $\lim z_n = z$ или $z_n \rightarrow z$, и говорят, что последовательность $\{z_n\}$ *сходится* к z . В противном случае последовательность $\{z_n\}$ называется *расходящейся*. Так же как и для последовательности действительных чисел, доказывается, что сходящаяся последовательность комплексных чисел не может иметь более одного предела. Это утверждение можно доказать, например, используя теорему, устанавливающую связь между сходимостью последовательности $\{z_n\}$ и сходимостью действительных последовательностей

$$\begin{aligned} x_1, x_2, \dots, x_n, \dots & \quad (x_n = \operatorname{Re} z_n) \\ y_1, y_2, \dots, y_n, \dots & \quad (y_n = \operatorname{Im} z_n)^1). \end{aligned}$$

Теорема. Для сходимости последовательности $\{z_n\}$ *необходимо и достаточно*, чтобы сходились последовательности $\{x_n\}$ и $\{y_n\}$.

Прежде всего заметим, что из определения модуля комплексного числа $z = x + iy$ вытекают следующие неравенства:

$$\begin{aligned} |x| &\leq |z| = \sqrt{x^2 + y^2} \leq |x| + |y| \\ |y| &\leq |z| = \sqrt{x^2 + y^2} \leq |x| + |y| \end{aligned} \quad (1)$$

которые в дальнейшем будут неоднократно использоваться.

Доказательство необходимости. Пусть $z = \lim z_n$. На основании (1) имеем:

$$|x - x_n| \leq |z - z_n|, \quad |y - y_n| \leq |z - z_n|.$$

Так как $z_n \rightarrow z$, то для любого $\varepsilon > 0$ найдется номер N , такой, что неравенство $|z - z_n| < \varepsilon$ будет выполняться для $n > N$, откуда получаем, что

$$|x - x_n| < \varepsilon, \quad |y - y_n| < \varepsilon \quad \text{для } n > N,$$

а это и означает сходимость последовательностей $\{x_n\}$ и $\{y_n\}$ соответственно к x и y .

Доказательство достаточности. Пусть $x_n \rightarrow x$, $y_n \rightarrow y$. Рассмотрим $z = x + iy$. На основании неравенства (1) запишем:

$$|z - z_n| \leq |x - x_n| + |y - y_n|.$$

По условию $x_n \rightarrow x$, $y_n \rightarrow y$, т. е. для любого числа $\varepsilon > 0$ существует номер N_1 , такой, что неравенство $|x - x_n| < \frac{\varepsilon}{2}$ выполняется при $n > N_1$ и номер N_2 , такой, что неравенство $|y - y_n| < \frac{\varepsilon}{2}$ выполняется при $n > N_2$. Тогда при $n > N$, где N — наибольшее из чисел N_1 и N_2 , выполняются оба неравенства и, следовательно,

$$|z - z_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{при } n > N,$$

т. е. последовательность $\{z_n\}$ сходится к z .

¹⁾ Здесь Re , или иногда пишут R , — начальные буквы слова *realis* — действительный (лат.), а Im , или иногда пишут I , начальные буквы слова *imaginaris* — мнимый (лат.).

Критерий Коши. Для сходимости последовательности $\{z_n\}$ необходимо и достаточно, чтобы для любого числа $\varepsilon > 0$ существовал такой номер N , что при $n > N$ и любом целом $p > 0$ выполнялось бы $|z_{n+p} - z_n| < \varepsilon$.

Доказательство получается непосредственно, если учесть только что доказанную теорему, критерий Коши для последовательности действительных чисел и воспользоваться неравенствами:

$$\begin{aligned} |x_{n+p} - x_n| &\leq |z_{n+p} - z_n| \leq |x_{n+p} - x_n| + |y_{n+p} - y_n|, \\ |y_{n+p} - y_n| &\leq |z_{n+p} - z_n| \leq |x_{n+p} - x_n| + |y_{n+p} - y_n|. \end{aligned}$$

Предлагается привести эти рассуждения самостоятельно.

Рассмотрим ряд $w_1 + w_2 + w_3 + \dots + w_n + \dots$ ($w_n = u_n + iv_n$) с комплексными членами.

Ряд $\sum_{n=1}^{\infty} w_n$ называется *сходящимся*, если сходится последовательность $\{S_n\}$, $S_n = \sum_{k=1}^n w_k$, частных сумм ряда.

Если $\lim S_n = S$, то число S называется *суммой ряда* и в этом случае записывают: $S = \sum_{n=1}^{\infty} w_n$. Если последовательность $\{S_n\}$ расходится, то ряд называется *расходящимся*.

Ряду $\sum w_n$ с комплексными членами соответствуют два действительных ряда: $\sum u_n$ ($u_n = \operatorname{Re} w_n$) и $\sum v_n$ ($v_n = \operatorname{Im} w_n$).

Теорема. Для сходимости ряда $\sum w_n$ необходимо и достаточно, чтобы сходились ряды $\sum u_n$ и $\sum v_n$.

Из критерия Коши для последовательности с комплексными членами получаем **критерий Коши для ряда с комплексными членами:**

Для сходимости ряда $\sum w_n$ необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовал такой номер N , что при $n > N$ и любом целом $p > 0$ выполнялось бы неравенство

$$|w_{n+1} + w_{n+2} + \dots + w_{n+p}| < \varepsilon.$$

Действительно, критерий Коши для ряда получается сразу же, если учесть, что

$$S_{n+p} - S_n = w_{n+1} + w_{n+2} + \dots + w_{n+p}.$$

Пусть $\sum |w_n|$ — ряд из модулей членов ряда $\sum w_n$. Так же как и для рядов с действительными членами, с помощью критерия Коши доказывается теорема: если ряд $\sum |w_n|$ сходится, то сходится и ряд $\sum w_n$.

Ряд $\sum w_n$ называют *абсолютно сходящимся*, если сходится ряд $\sum |w_n|$, если же ряд $\sum |w_n|$ расходится, а ряд $\sum w_n$ сходится, то ряд $\sum w_n$ называется *неабсолютно сходящимся*.

Легко видеть, что к рядам с комплексными членами применимы признаки сходимости Даламбера и Коши, сформулированные на стр. 599.

Пример. Ряд $\sum_{n=1}^{\infty} \frac{(1+i)^n}{n!}$ сходится абсолютно, так как, применяя признак Даламбера к ряду из модулей его членов, получим:

$$\lim_{n \rightarrow \infty} \frac{|u_{n+1}|}{|u_n|} = \lim_{n \rightarrow \infty} \frac{|1+i|^{n+1} n!}{(n+1)! |1+i|^n} = \lim_{n \rightarrow \infty} \frac{|1+i|}{n+1} = \lim_{n \rightarrow \infty} \frac{\sqrt{2}}{n+1} = 0 < 1,$$

Теорема. Для абсолютной сходимости ряда $\sum w_n$ необходима и достаточно абсолютная сходимость рядов $\sum u_n$ и $\sum v_n$.

Легко доказать эту теорему, воспользовавшись неравенствами

$$\begin{aligned} |u_n| &\leq |w_n| \leq |u_n| + |v_n|, \\ |v_n| &\leq |w_n| \leq |u_n| + |v_n| \end{aligned}$$

и признаком сравнения для рядов с неотрицательными членами.

Из этой теоремы, а также из теоремы о перестановке членов абсолютно сходящихся рядов с действительными членами получаем теорему: в абсолютно сходящемся ряде $\sum w_n$ любая перестановка членов ряда сохраняет абсолютную сходимость ряда и величину его суммы. Также можно показать, что в сходящемся ряде $\sum w_n$ любая группировка членов ряда, не изменяющая их порядка, сохраняет сходимость ряда и величину его суммы. В абсолютно сходящемся ряде $\sum w_n$ допустима любая перестановка и группировка членов ряда.

Не доказывая, заметим, что теоремы об умножении сходящегося ряда на число, в общем случае комплексное, о сложении сходящихся рядов и теорема об умножении абсолютно сходящихся рядов остаются справедливыми и для рядов с комплексными членами.

Раздел 2

ФУНКЦИОНАЛЬНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ И РЯДЫ

§ 6. Сходимость и равномерная сходимость

Рассмотрим бесконечную последовательность функций

$$f_1(x), f_2(x), f_3(x), \dots, f_n(x), \dots, \quad (1)$$

непрерывных на некотором промежутке.

Положив $x = x_0$ (x_0 из рассматриваемого промежутка), получим числовую последовательность $f_1(x_0), f_2(x_0), \dots, f_n(x_0), \dots$, которая может сходиться, а может расходиться. Если эта последовательность сходится, то говорят, что *функциональная последовательность (1) сходится* при $x = x_0$ (или в точке x_0). Пределом этой последовательности будет число, зависящее от выбранного значения $x = x$. При одних значениях x последовательность (1) может сходиться, а при других — расходиться. Совокупность всех значений x , для которых последовательность (1) сходится, называется *областью сходимости* этой последовательности.

Так как для каждого значения x из области сходимости последовательности пределом будет число, то в области сходимости последовательности $\{f_n(x)\}$ определяется функция $f(x)$ — предел функциональной последовательности. В этом случае говорят, что функциональная последовательность $\{f_n(x)\}$ или, что то же, последовательность функций $\{f_n(x)\}$, сходится к функции $f(x)$ в соответствующей области и записывают так:

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad \text{или} \quad f_n(x) \rightarrow f(x).$$

Пример 1. Рассмотрим последовательность

$$1, x, x^2, \dots, x^{n-1}, \dots$$

$$\lim_{n \rightarrow \infty} x^{n-1} = \begin{cases} 0 & \text{при } |x| < 1, \\ 1 & \text{при } x = 1. \end{cases}$$

При $x = -1$ и $|x| > 1$ последовательность расходится, следовательно, областью сходимости является полуинтервал $(-1, 1]$. Предельная функция:

$$f(x) = \begin{cases} 0 & \text{при } |x| < 1, \\ 1 & \text{при } x = 1. \end{cases}$$

Пример 2. Последовательность $\{f_n(x)\} = \{\cos^n x\}$ сходится при всех значениях x , кроме $x = (2k+1)\pi$ ($k=0, \pm 1, \pm 2, \dots$), так как

$$\lim_{n \rightarrow \infty} \cos^n x = \begin{cases} 0 & \text{при } x \neq k\pi \text{ (} k \text{ — целое)}, \\ 1 & \text{при } x = 2k\pi \text{ (} k \text{ — целое)}. \end{cases}$$

Пример 3. Для последовательности функций e^{-x} , e^{-2x} , ..., e^{-nx} , ..., предел которой

$$\lim_{n \rightarrow \infty} e^{-nx} = \begin{cases} 0 & \text{при } x > 0, \\ 1 & \text{при } x = 0, \end{cases}$$

областью сходимости является совокупность неотрицательных значений x , $x \geq 0$.

Пример 4. Последовательность $\{f_n(x)\} = \{n^n x^n\}$ сходится только в одной точке $x=0$ и ее предел в этой точке равен нулю, так как при любом x , отличном от нуля, начиная с некоторого n , выполняется неравенство $|n^n x^n| > 1$.

Ради простоты будем рассматривать сходимость функциональной последовательности на отрезке, хотя все понятия и теоремы этого параграфа переносятся на любой промежуток.

Учитывая определение предела числовой последовательности, можно следующим образом определить сходимость функциональной последовательности.

Последовательность $\{f_n(x)\}$ сходится к $f(x)$ на отрезке $[a, b]$, если для любого числа $\epsilon (\epsilon > 0)$ и любой точки x из рассматриваемого отрезка существует номер $N = N(\epsilon, x)$, такой, что неравенство

$$|f(x) - f_n(x)| < \epsilon$$

выполняется при $n > N$. Заметим, что при заданном $\epsilon > 0$ номер N зависит от ϵ и рассматриваемого значения x отрезка $[a, b]$.

При выбранном $\epsilon > 0$ каждой точке отрезка $[a, b]$ соответствует свой номер N , следовательно, номеров, соответствующих всем точкам отрезка $[a, b]$, будет, вообще говоря, бесчисленное множество. Если из всех этих номеров можно выбрать наибольший (как известно, это можно сделать не всегда, если

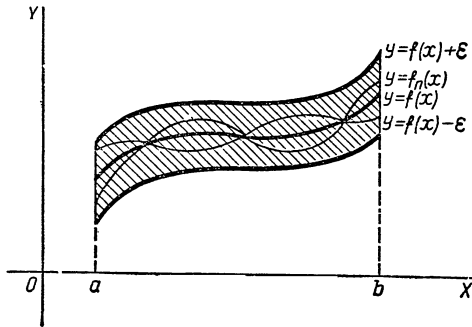


Рис. 433.

чисел N бесконечное множество), то этот номер будет «годиться» для всех точек отрезка $[a, b]$, т. е. будет общим для всех точек этого отрезка.

Если для каждого $\epsilon > 0$ можно выбрать номер N , «годный» для всех точек отрезка $[a, b]$, то мы имеем так называемую равномерную сходимость последовательности на отрезке $[a, b]$. Итак, последовательность $\{f_n(x)\}$ равномерно сходится к $f(x)$ на отрезке $[a, b]$, если для любого числа $\epsilon > 0$ найдется номер $N = N(\epsilon)$, такой, что неравенство $|f(x) - f_n(x)| < \epsilon$ выполняется при $n > N$ для всех точек отрезка $[a, b]$.

Таким образом, для всех x отрезка $[a, b]$ выполняются неравенства

$$f(x) - \epsilon < f_n(x) < f(x) + \epsilon \quad (2)$$

при $n > N(\epsilon)$. Пусть ϵ — какое-нибудь положительное число; « ϵ -полоской» графика функции $f(x)$ на отрезке $[a, b]$ называется совокупность точек (x, y) плоскости, координаты которых удовлетворяют условиям:

$$a \leq x \leq b, \quad f(x) - \epsilon < y < f(x) + \epsilon.$$

Если учесть неравенства (2), то равномерная сходимость последовательности $\{f_n(x)\}$ к функции $f(x)$ на отрезке $[a, b]$ геометрически означает: какую бы « ϵ -полоску» графика функции $f(x)$ ни взяли, графики функций $f_n(x)$ при $n > N$ на отрезке $[a, b]$ не выходят из этой полоски (рис. 433).

Рассмотрим примеры.

Пример 1. Последовательность

$$\frac{\sin x}{1}, \frac{\sin 2x}{2}, \dots, \frac{\sin nx}{n}, \dots \quad (3)$$

сходится на всей числовой оси к функции $f(x) \equiv 0$, так как

$$\lim_{n \rightarrow \infty} \frac{\sin nx}{n} = 0, \quad -\infty < x < +\infty.$$

Графики функций этой последовательности изображены на рис. 434.

Сходимость последовательности (3) — равномерная на любом отрезке $[a, b]$. Действительно, рассмотрим какой-нибудь отрезок $[a, b]$ и возьмем ε -полоску графика функции $f(x) \equiv 0$ на этом отрезке. Нетрудно видеть, что графики всех членов последовательности, кроме, возможно, конечного числа их, на отрезке $[a, b]$ попадут в эту полоску. Равномерную сходимость последовательности (3) на отрезке $[a, b]$ можно доказать и аналитически. В самом деле, возьмем любое число $\varepsilon > 0$ и определим n

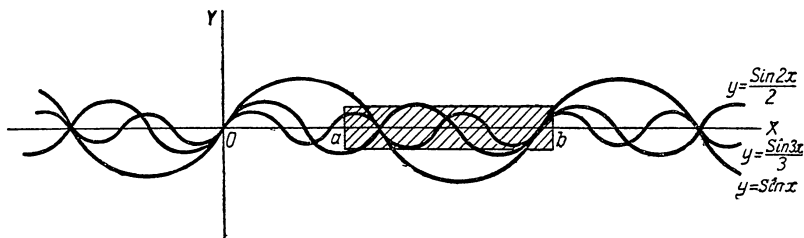


Рис. 434.

из неравенства $\left|0 - \frac{\sin nx}{n}\right| < \varepsilon$. Так как $\left|0 - \frac{\sin nx}{n}\right| \leq \frac{1}{n}$, то при n , удовлетворяющих усиленному неравенству $\frac{1}{n} < \varepsilon$, $n > \frac{1}{\varepsilon}$, выполняется и исходное неравенство. Поэтому, для заданного $\varepsilon > 0$ неравенство $\left|0 - \frac{\sin nx}{n}\right| < \varepsilon$ выполняется для всех точек отрезка $[a, b]$ при $n > N$, где N — натуральное число, не меньшее $\frac{1}{\varepsilon}$, т. е. N зависит только от ε .

Последовательность $\left\{\frac{\sin nx}{n}\right\}$ сходится равномерно и на промежутке $-\infty < x < +\infty$.

Пример 2. Рассмотрим последовательность функций $x, x^2, x^3, \dots, x^n, \dots$ на отрезке $[0, 1]$. В этом случае

$$f(x) = \lim_{n \rightarrow \infty} x^n = \begin{cases} 0 & \text{при } 0 \leq x < 1, \\ 1 & \text{при } x = 1. \end{cases}$$

Последовательность $\{x^n\}$ на отрезке $[0, 1]$ сходится неравномерно, так как, какую бы ε -полоску ($0 < \varepsilon < 1$) графика функции $f(x)$ мы ни взяли, графики всех функций последовательности на отрезке $[0, 1]$ выходят из нее (рис. 435). Но сходимость будет равномерной на любом отрезке $[0, a]$ при $a < 1$. Действительно в этом случае для любого $\varepsilon > 0$ графики всех функций последовательности

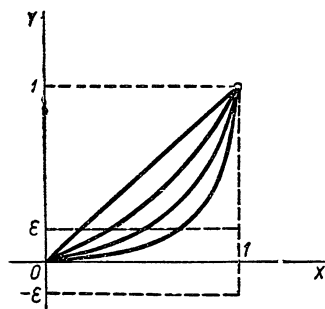


Рис. 435.

$\{x^n\}$, кроме конечного числа их, находятся внутри ε -полоски графика функции $f(x) \equiv 0$. Здесь только лишь конечное число графиков функций последовательности $\{x^n\}$ пересекает прямую $y = \varepsilon$ на отрезке $[0, a]$ (все остальные графики пересекают прямую $y = \varepsilon$ вне отрезка $[0, a]$).

Теперь этот же результат получим аналитически. Из определения равномерной сходимости последовательности непосредственно следует: если последовательность равномерно сходится на отрезке $[a, b]$, то она равномерно сходится на любом промежутке, принадлежащем этому отрезку, в частности на интервале (a, b) и на каждом полуинтервале с концами a и b . Поэтому достаточно показать, что последовательность $\{x^n\}$ сходится неравномерно на полуинтервале $[0, 1)$.

Возьмем любое число $\varepsilon > 0$ и рассмотрим неравенство $|0 - x^n| < \varepsilon$ или, что то же, $x^n < \varepsilon$, так как $0 \leq x < 1$. Если $x = 0$, то неравенство

$x^n < \epsilon$ выполняется при всех n , если $0 < x < 1$, то, взяв обратное неравенство ¹⁾ $\frac{1}{x^n} > \frac{1}{\epsilon}$ и прологарифмировав его, получим:

$$n \ln \frac{1}{x} > \ln \frac{1}{\epsilon},$$

откуда определится

$$n > \frac{\ln \frac{1}{\epsilon}}{\ln \frac{1}{x}}.$$

Следовательно, если за N принять натуральное число, не меньшее $\frac{\ln \frac{1}{\epsilon}}{\ln \frac{1}{x}}$ (N зависит от ϵ и x), то в каждой точке интервала $(0, 1)$ при $n > N$

будет выполняться неравенство $|0 - x^n| < \epsilon$.

Здесь N является функцией ϵ и x , $N = N(\epsilon, x)$, при $x \rightarrow 1$ $N \rightarrow \infty$ и общего N для всех точек полуинтервала $[0, 1)$ (или отрезка $[0, 1]$) выбрать нельзя. Таким образом, сходимость последовательности неравномерная на отрезке $[0, 1]$.

Нетрудно видеть, что на отрезке $[0, a]$, при $a < 1$, последовательности $\{x^n\}$ сходится равномерно, в этом случае оказывается можно выбрать номер N , общий для всех точек отрезка $[0, a]$, достаточно за N принять

натуральное число, не меньшее $\frac{\ln \frac{1}{\epsilon}}{\ln \frac{1}{a}}$,

Критерий Коши равномерной сходимости последовательности. Для того чтобы последовательность функций $f_1(x), f_2(x), \dots, f_n(x), \dots$ была равномерно сходящейся на отрезке $[a, b]$, необходимо и достаточно, чтобы для любого числа $\epsilon > 0$ существовал такой номер $N = N(\epsilon)$, зависящий только от ϵ , что при $n > N$ и любом целом $p > 0$ неравенство

$$|f_{n+p}(x) - f_n(x)| < \epsilon$$

выполнялось бы для всех x отрезка $[a, b]$.

Доказательство необходимости. По условию $f_n(x) \rightarrow f(x)$ равномерно на отрезке $[a, b]$, следовательно, для любого числа $\epsilon > 0$ существует номер $N = N(\epsilon)$, такой, что при $n > N$ и всех x отрезка $[a, b]$ выполняется неравенство

$$|f(x) - f_n(x)| < \frac{\epsilon}{2}.$$

Если $n > N$ и p — любое целое положительное число, то неравенство $|f(x) - f_{n+p}(x)| < \frac{\epsilon}{2}$ также выполняется для всех x отрезка $[a, b]$.

Следовательно, для всех x отрезка $[a, b]$ при $n > N$ и любом целом положительном p , имеем:

$$\begin{aligned} |f_{n+p}(x) - f_n(x)| &= |[f_{n+p}(x) - f(x)] + [f(x) - f_n(x)]| \leq \\ &\leq |f_{n+p}(x) - f(x)| + |f(x) - f_n(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \end{aligned}$$

т. е. $|f_{n+p}(x) - f_n(x)| < \epsilon$.

Доказательство достаточности. По условию для любого числа $\epsilon > 0$ найдется номер N , такой, что при $n > N$ и целом $p > 0$ для всех x отрезка $[a, b]$ выполняется неравенство

$$|f_{n+p}(x) - f_n(x)| < \frac{\epsilon}{2}. \quad (4)$$

¹⁾ Обратное неравенство рассматривается затем, чтобы в результате логарифмирования неравенства не иметь дело с отрицательными значениями логарифмов.

Следовательно, условие (4) выполняется для каждого фиксированного значения x отрезка $[a, b]$. Но при фиксированном x условие (4) является условием критерия Коши сходимости числовой последовательности, соответствующей рассматриваемому значению x . Учитывая это, получим, что последовательность $\{f_n(x)\}$ сходится в любой точке отрезка $[a, b]$, т. е. она сходится на отрезке $[a, b]$. Пусть на этом отрезке $\lim_{n \rightarrow \infty} f_n(x) = f(x)$.

Чтобы доказать равномерную сходимость последовательности $\{f_n(x)\}$ на этом отрезке $[a, b]$, перейдем к пределу при $p \rightarrow \infty$ в неравенстве (4), которое выполняется при всех x отрезка $[a, b]$.

Тогда получим неравенство $|f(x) - f_n(x)| \leq \frac{\varepsilon}{2} < \varepsilon$, справедливое при $n > N$ и всех x отрезка $[a, b]$, а это означает, что последовательность $\{f_n(x)\}$ сходится к функции $f(x)$ равномерно на отрезке $[a, b]$.

Функциональные ряды. Функциональным рядом называется ряд

$$u_1(x) + u_2(x) + \dots + u_n(x) + \dots$$

или, что то же, $\sum_{n=1}^{\infty} u_n(x)$, членами которого являются функции от x , определенные на некотором множестве. Функция $u_n(x)$ называется *общим членом ряда*.

Частными (или *частичными*) *суммами ряда* $\sum_{n=1}^{\infty} u_n(x)$ называются

функции $S_n(x) = \sum_{k=1}^n u_k(x) \quad (n = 1, 2, \dots)$.

Функциональный ряд $\sum_{n=1}^{\infty} u_n(x)$ называется *сходящимся при $x = x_0$*

(в точке x_0), если в этой точке сходится последовательность его частных сумм:

$$S_1(x_0), S_2(x_0), \dots, S_n(x_0), \dots$$

Или, другими словами, функциональный ряд $\sum_{n=1}^{\infty} u_n(x)$ сходится при

$x = x_0$, если сходится числовой ряд $\sum_{n=1}^{\infty} u_n(x_0)$. Предел последовательности

$\{S_n(x_0)\}$, обозначим его $f(x_0)$, называется суммой ряда $\sum_{n=1}^{\infty} u_n(x)$

в точке x_0 . Совокупность всех значений x , для которых сходится ряд

$\sum_{n=1}^{\infty} u_n(x)$, называется *областью сходимости* этого ряда.

Учитывая определение сходимости ряда в точке, получим, что область сходимости ряда $\sum_{n=1}^{\infty} u_n(x)$ является областью сходимости последовательности $\{S_n(x)\}$ его частных сумм; сумма ряда (определенная в области его сходимости) является пределом последовательности частных сумм этого ряда.

Как и в случае последовательности, сходимость функционального ряда будем рассматривать на отрезке, помня, что в этом параграфе слово «отрезок» можно заменить словом «промежуток».

Пусть $\lim_{n \rightarrow \infty} S_n(x) = f(x)$ на отрезке $[a, b]$, тогда $f(x) = \sum_{n=1}^{\infty} u_n(x)$ на рассматриваемом отрезке. В этом случае говорят также, что функция $f(x)$ *разлагается в ряд на отрезке $[a, b]$* .

Как было сказано, сходимость функционального ряда на отрезке $[a, b]$ означает, что для любого значения x отрезка $[a, b]$ соответствующий числовой ряд сходится. Поэтому для исследования на сходимость функциональных рядов можно использовать признаки сходимости числовых рядов.

Рассмотрим примеры.

Пример 1. Ряд $\sum_{n=1}^{\infty} \frac{\sin nx}{n^2}$ сходится при всех значениях x , так как неравенство $\left| \frac{\sin nx}{n^2} \right| \leq \frac{1}{n^2}$ выполняется при любом значении x и ряд $\sum_{n=1}^{\infty} \frac{1}{n^2}$ сходится. Следовательно, по признаку сравнения рядов данный ряд сходится на всей числовой оси.

Пример 2. Рассмотрим ряд $\sum_{n=1}^{\infty} \frac{x^n}{n^3}$.

Применяя признак Даламбера к ряду из модулей его членов, получим:

$$\lim_{n \rightarrow \infty} \frac{|x|^{n+1} n^3}{(n+1)^3 |x|^n} = |x|.$$

Следовательно, при $|x| < 1$ ряд сходится, при $|x| > 1$ ряд расходится.

Остается рассмотреть $x=1$ и $x=-1$, при этих значениях получаем соответственно ряды $\sum_{n=1}^{\infty} \frac{1}{n^3}$ и $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^3}$, которые, как известно, сходятся. Окончательно получаем, что ряд $\sum_{n=1}^{\infty} \frac{x^n}{n^3}$ сходится при $|x| \leq 1$.

Пример 3. Как известно, ряд геометрической прогрессии $1 + x + x^2 + \dots + x^{n-1} + \dots$ сходится только при $|x| < 1$ и его сумма равна $\frac{1}{1-x}$, т. е.

$$\frac{1}{1-x} = 1 + x + x^2 + \dots + x^{n-1} + \dots$$

при $|x| < 1$.

Учитывая пример 3, получим, что ряд $\sum_{n=1}^{\infty} e^{nx^2}$ всюду расходится, так как это ряд геометрической прогрессии со знаменателем e^{x^2} .

Ряд $\sum_{n=1}^{\infty} \frac{1}{x^n}$ — также ряд геометрической прогрессии, он сходится при x , удовлетворяющих условию $\left| \frac{1}{x} \right| < 1$, т. е. когда $|x| > 1$.

Область сходимости этого ряда состоит из двух интервалов: $(-\infty, -1)$ и $(1, +\infty)$.

Пример 4. Ряд $\sum_{n=1}^{\infty} \frac{1}{n^x}$ сходится при $x > 1$ и расходится при $x \leq 1$. Область сходимости этого ряда — промежуток $(1, +\infty)$.

Ряд $\sum_{n=1}^{\infty} u_n(x)$ называется *равномерно сходящимся* на отрезке $[a, b]$, если равномерно сходится на этом отрезке последовательность частных сумм этого ряда. Применяя критерий Коши равномерной сходимости последовательности к последовательности частных сумм ряда, получим критерий Коши равномерной сходимости ряда: для равномерной сходимости ряда $\sum_{n=1}^{\infty} u_n(x)$ на отрезке $[a, b]$ необходимо и достаточно, чтобы для любого числа $\epsilon > 0$ существовал такой номер $N(\epsilon)$, что при $n > N$ и

любом целом $p > 0$ неравенство $|u_{n+1}(x) + u_{n+2}(x) + \dots + u_{n+p}(x)| < \varepsilon$ выполнялось бы для всех x отрезка $[a, b]$.

Действительно, это получается сразу же, если учесть, что

$$S_{n+p}(x) - S_n(x) = u_{n+1}(x) + u_{n+2}(x) + \dots + u_{n+p}(x).$$

Следует иметь в виду, что выделение из рядов сходящихся рядов равномерно сходящихся, как это будет видно дальше, является существенным, так как равномерно сходящиеся ряды «во многих отношениях» ведут себя так же, как конечные суммы, тогда как о рядах, сходящихся неравномерно, этого сказать нельзя.

Практическое применение критерия Коши для исследования ряда на равномерную сходимость затруднительно, поэтому рассмотрим простой и удобный, в смысле практического применения, достаточный признак равномерной сходимости ряда — признак Вейерштрасса¹⁾.

Сначала введем определение. Говорят, что функциональный ряд $\sum_{n=1}^{\infty} u_n(x)$ мажорируется на отрезке $[a, b]$ числовым рядом $\sum_{n=1}^{\infty} a_n$, если неравенство $|u_n(x)| \leq a_n$ выполняется при любом n и всех x отрезка $[a, b]$.

Признак равномерной сходимости Вейерштрасса. Если ряд $\sum_{n=1}^{\infty} u_n(x)$ на отрезке $[a, b]$ мажорируется числовым сходящимся рядом, то он равномерно сходится на этом отрезке²⁾.

Доказательство. По условию ряд $\sum_{n=1}^{\infty} a_n$ сходится, поэтому на основании критерия Коши для числовых рядов для любого $\varepsilon > 0$ существует номер N , такой, что при $n > N$ и любом целом $p > 0$ выполняется неравенство $|a_{n+1} + a_{n+2} + \dots + a_{n+p}| < \varepsilon$, т. е., так как $a_n \geq 0$, неравенство

$$a_{n+1} + a_{n+2} + \dots + a_{n+p} < \varepsilon.$$

При $n > N$ и всех целых $p > 0$

$$\begin{aligned} & |u_{n+1}(x) + u_{n+2}(x) + \dots + u_{n+p}(x)| \leq \\ & \leq |u_{n+1}(x)| + |u_{n+2}(x)| + \dots + |u_{n+p}(x)| \leq a_{n+1} + a_{n+2} + \dots + a_{n+p} < \varepsilon, \end{aligned}$$

для всех x отрезка $[a, b]$. А значит, по критерию Коши равномерной сходимости ряда, ряд $\sum_{n=1}^{\infty} u_n(x)$ равномерно сходится на отрезке $[a, b]$, что и требовалось доказать.

Нетрудно видеть, что если ряд $\sum_{n=1}^{\infty} u_n(x)$ на отрезке $[a, b]$ мажорируется числовым сходящимся рядом, то он сходится на этом отрезке абсолютно.

Пример 1. Ряды $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$ и $\sum_{n=1}^{\infty} \frac{1}{n^2 + x^2}$ равномерно сходятся на всей числовой оси, так как для всех x , $-\infty < x < +\infty$, $\left| \frac{\cos nx}{n^2} \right| \leq \frac{1}{n^2}$ и $\left| \frac{1}{n^2 + x^2} \right| \leq \frac{1}{n^2}$, а ряд $\sum_{n=1}^{\infty} \frac{1}{n^2}$ сходится.

¹⁾ Вейерштрасс (1815—1897) — немецкий математик.

²⁾ В формулировке этой теоремы рассматривается отрезок $[a, b]$, но как уже было замечено, все рассуждения этого параграфа, и в частности признак Вейерштрасса, остаются верными, если вместо отрезка рассматривать любой промежуток, как конечный, так и бесконечный.

Пример 2. Ряд $\sum_{n=1}^{\infty} \frac{x^n}{n^3}$ равномерно сходится на отрезке $[-1, 1]$, так как $\left| \frac{x^n}{n^3} \right| \leq \frac{1}{n^3}$ при $|x| \leq 1$. При $|x| > 1$, как было показано, см. пример 2 на стр. 610, ряд $\sum_{n=1}^{\infty} \frac{x^n}{n^3}$ расходится.

Свойства равномерно сходящихся последовательностей и рядов

Теорема о непрерывности предельной функции. Если последовательность функций, непрерывных на отрезке $[a, b]$

$$f_1(x), f_2(x), \dots, f_n(x), \dots,$$

равномерно сходится к функции $f(x)$ на этом отрезке, то $f(x)$ — непрерывная функция на отрезке $[a, b]$.

Доказательство. Возьмем произвольную точку x_0 интервала (a, b) и докажем непрерывность функции $f(x)$ в этой точке.

По условию теоремы последовательность $\{f_n(x)\}$ равномерно сходится к функции $f(x)$ на отрезке $[a, b]$, следовательно, для любого выбранного числа $\varepsilon > 0$ найдется число n , такое, что для всех точек x интервала (a, b) выполняется неравенство

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3}.$$

В частности, оно выполняется и в точке x_0 и в любой точке $x_0 + \Delta x$, принадлежащей интервалу (a, b) .

Так как $f_n(x)$ — непрерывная функция на отрезке $[a, b]$, а следовательно, и в точке x_0 , то для выбранного числа ε найдется такое число $\delta > 0$, что при $|\Delta x| < \delta$ справедливо следующее неравенство

$$|f_n(x_0 + \Delta x) - f_n(x_0)| < \frac{\varepsilon}{3}.$$

При этом

$$\begin{aligned} |f(x_0 + \Delta x) - f(x_0)| &= |[f(x_0 + \Delta x) - f_n(x_0 + \Delta x)] + [f_n(x_0 + \Delta x) - f_n(x_0)] + \\ &\quad + [f_n(x_0) - f(x_0)]| \leq |f(x_0 + \Delta x) - f_n(x_0 + \Delta x)| + \\ &\quad + |f_n(x_0 + \Delta x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

т. е. $|f(x_0 + \Delta x) - f(x_0)| < \varepsilon$ при $|\Delta x| < \delta$.

Так как $\varepsilon > 0$ выбиралось произвольно, то на основании полученного неравенства $|f(x_0 + \Delta x) - f(x_0)| < \varepsilon$, справедливого при $|\Delta x| < \delta$, заключаем, что $f(x)$ — непрерывная функция в точке x_0 .

Учитывая, что x_0 — любая точка интервала (a, b) , получаем непрерывность функции $f(x)$ на интервале (a, b) . Так как условия теоремы выполняются на отрезке $[a, b]$, то нужно доказать еще непрерывность функции $f(x)$ справа в точке $x = a$ и слева в точке $x = b$.

Односторонняя непрерывность функции $f(x)$ на концах отрезка $[a, b]$ устанавливается так же, как и во внутренней точке этого отрезка, с той лишь разницей, что при доказательстве непрерывности справа в точке a рассматриваются $\Delta x > 0$, а при доказательстве непрерывности слева в точке b рассматриваются $\Delta x < 0$.

Таким образом, теорема полностью доказана.

Следует заметить, что условие равномерной сходимости в формулировке теоремы опустить нельзя, так как последовательность непрерывных на отрезке функций может сходиться к разрывной предельной функции на этом отрезке (см. $\{x^n\}$ на $[0, 1]$, рис. 435). Из того, что предельная

функция последовательности непрерывных функций непрерывна на некотором отрезке, не следует равномерной сходимости этой последовательности на рассматриваемом отрезке, т. е. условие равномерной сходимости последовательности является только достаточным для непрерывности предельной функции.

В самом деле, рассмотрим последовательность функций $\{f_n(x)\}$, определенных на отрезке $[0, 1]$ следующим образом:

$$f_n(x) = \begin{cases} 0 & \text{при } x=0 \text{ и } \frac{1}{2^{n-1}} \leq x \leq 1, \\ 2^n & \text{при } x = \frac{1}{2^n}, \\ \text{линейная} & \text{на интервалах } \left(0, \frac{1}{2^n}\right) \text{ и } \left(\frac{1}{2^n}, \frac{1}{2^{n-1}}\right). \end{cases}$$

Графики таких функций изображены на рис. 436. Последовательность $\{f_n(x)\}$ сходится на отрезке $[0, 1]$ к функции $f(x) \equiv 0$ неравномерно (так как, какую бы ε -полоску ($0 < \varepsilon < 1$) графика функции $f(x) \equiv 0$ на отрезке $[0, 1]$ мы ни взяли, графики всех функций последовательности выходят из нее), хотя предельная функция непрерывна на отрезке $[0, 1]$.

Как следствие теоремы о непрерывности предельной функции последовательности получается аналогичная теорема для ряда.

Теорема. Сумма равномерно сходящегося на отрезке $[a, b]$ ряда $\sum_{n=1}^{\infty} u_n(x)$ непрерывных функций есть непрерывная функция на отрезке $[a, b]$.

Действительно, последовательность $\{S_n(x)\}$ частных сумм ряда является последовательностью непрерывных функций на отрезке $[a, b]$, так как сумма конечного числа непрерывных функций на отрезке — непрерывная функция на этом отрезке.

Кроме того, последовательность $\{S_n(x)\}$ равномерно сходится на отрезке $[a, b]$. Поэтому на основании доказанной теоремы для последовательности заключаем, что сумма ряда — непрерывная функция на отрезке $[a, b]$.

Заметим, что это свойство равномерно сходящихся рядов аналогично следующему свойству конечных сумм: сумма конечного числа непрерывных на отрезке функций — непрерывная функция на этом отрезке.

Рассмотрим теперь теоремы о почленном интегрировании и дифференцировании рядов, отметив их особую значимость для приложений.

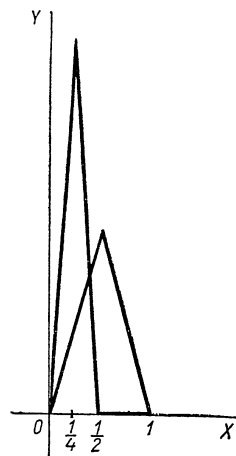


Рис. 436.

§ 7. Интегрирование и дифференцирование функциональных последовательностей и рядов

Теорема. Если последовательность непрерывных функций

$$f_1(x), f_2(x), \dots, f_n(x), \dots$$

на отрезке $[a, b]$ равномерно сходится к функции $f(x)$ на этом отрезке, то числовая последовательность

$$\int_a^b f_1(x) dx, \int_a^b f_2(x) dx, \dots, \int_a^b f_n(x) dx, \dots$$

сходится и ее предел равен $\int_a^b f(x) dx$:

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx$$

или, что то же,

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx.$$

Таким образом, если условия теоремы выполняются, то знаки предела и интеграла можно переставлять.

Доказательство. Так как все функции последовательности $\{f_n(x)\}$ непрерывны на отрезке и последовательность $\{f_n(x)\}$ равномерно сходится на этом отрезке, то по доказанной теореме $f(x)$ — непрерывная функция на отрезке $[a, b]$. А, как известно, интегралы от непрерывных функций существуют.

$$\text{Теперь докажем, что } \int_a^b f_n(x) dx \rightarrow \int_a^b f(x) dx.$$

Действительно, по условию $f_n(x) \rightarrow f(x)$ равномерно на отрезке $[a, b]$, а это значит, что для любого числа $\varepsilon > 0$ существует номер N , такой, что при $n > N$ для всех точек x отрезка $[a, b]$ выполняется неравенство $|f(x) - f_n(x)| < \frac{\varepsilon}{b-a}$.

Пусть n — любое натуральное число, удовлетворяющее условию $n > N$. Тогда модуль разности интегралов оценивается следующим образом:

$$\begin{aligned} \left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| &= \left| \int_a^b [f(x) - f_n(x)] dx \right| \leq \\ &\leq \int_a^b |f(x) - f_n(x)| dx \leq \frac{\varepsilon}{(b-a)} \cdot (b-a) = \varepsilon. \end{aligned}$$

Таким образом,

$$\left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| < \varepsilon \text{ при } n > N.$$

Отсюда получаем:

$$\int_a^b f_n(x) dx \rightarrow \int_a^b f(x) dx.$$

Заметим, что в случае неравномерной сходимости последовательности теорема может быть неверна, как это видно на примере последовательности $\{f_n(x)\}_1$ (рис. 326). Действительно, в этом случае при любом n $\int_0^1 f_n(x) dx = 1$, но $\int_0^1 f(x) dx = 0$ и, следовательно, предел последовательности $\{\int_0^1 f_n(x) dx\}$ не равен интегралу $\int_0^1 f(x) dx$.

Как следствие доказанной теоремы для последовательности получается следующая теорема для ряда.

Теорема. Равномерно сходящийся на отрезке $[a, b]$ ряд с непрерывными членами можно почленно интегрировать на этом отрезке, т. е. ряд, составленный из интегралов от его членов по отрезку $[a, b]$, сходится к интегралу от суммы ряда по этому отрезку.

Действительно, пусть $f(x) = u_1(x) + u_2(x) + \dots + u_n(x) + \dots$ на отрезке $[a, b]$, причем члены ряда непрерывны и ряд сходится равномерно на этом отрезке. Тогда $S_n(x)$ — непрерывная функция на отрезке $[a, b]$, $S_n(x) \rightarrow f(x)$ равномерно на отрезке $[a, b]$, и по доказанной теореме для последовательности имеем:

$$\lim_{n \rightarrow \infty} \int_a^b S_n(x) dx = \int_a^b f(x) dx. \quad (1)$$

Но для конечного числа слагаемых интеграл от суммы функций равен сумме от интегралов слагаемых:

$$\begin{aligned} \int_a^b S_n(x) dx &= \int_a^b [u_1(x) + u_2(x) + \dots + u_n(x)] dx = \\ &= \int_a^b u_1(x) dx + \int_a^b u_2(x) dx + \dots + \int_a^b u_n(x) dx. \end{aligned}$$

При этом равенство (1) запишется так:

$$\lim_{n \rightarrow \infty} \left[\int_a^b u_1(x) dx + \int_a^b u_2(x) dx + \dots + \int_a^b u_n(x) dx \right] = \int_a^b f(x) dx$$

или так:

$$\int_a^b u_1(x) dx + \int_a^b u_2(x) dx + \dots + \int_a^b u_n(x) dx + \dots = \int_a^b f(x) dx,$$

т. е. сумма ряда $\sum_{n=1}^{\infty} \int_a^b u_n(x) dx$ равна $\int_a^b f(x) dx$ или, что то же,

$$\sum_{n=1}^{\infty} \int_a^b u_n(x) dx = \int_a^b \sum_{n=1}^{\infty} u_n(x) dx.$$

Заметим, что полученное свойство аналогично свойству конечной суммы: интеграл от суммы функций равен сумме интегралов от слагаемых.

Пример. Ряд $1 + x + x^2 + \dots + x^n + \dots$ сходится равномерно на отрезке $\left[-\frac{1}{2}, \frac{1}{2}\right]$, так как $|x^n| \leq \frac{1}{2^n}$, т. е. данный ряд мажорируется на отрезке $\left[-\frac{1}{2}, \frac{1}{2}\right]$

числовым сходящимся рядом $\sum_{n=0}^{\infty} \frac{1}{2^n}$. Как известно,

$$1 + x + x^2 + \dots + x^n + \dots = \frac{1}{1-x} \text{ при } |x| < 1.$$

Интегрируя этот ряд на отрезке $\left[0, \frac{1}{2}\right]$, получим:

$$\int_0^{\frac{1}{2}} \frac{dx}{1-x} = \int_0^{\frac{1}{2}} (1 + x + x^2 + \dots + x^n + \dots) dx$$

или

$$-\ln(1-x) \Big|_0^{\frac{1}{2}} = \ln 2 = \frac{1}{2} + \frac{1}{2 \cdot 2^2} + \frac{1}{3 \cdot 2^3} + \dots + \frac{1}{n \cdot 2^n} + \dots$$

Перейдем теперь к дифференцированию последовательностей.

Прежде всего заметим, что если последовательность дифференцируемых функций $f_1(x), f_2(x), \dots, f_n(x), \dots$ равномерно сходится к функ-

ции $f(x)$ на отрезке $[a, b]$, то последовательность производных этих функций: $f'_1(x), f'_2(x), \dots, f'_n(x), \dots$ может не сходиться к функции $f'(x)$ на этом отрезке. Это подтверждается примером последовательности функций:

$$\sin x, \frac{\sin 2x}{2}, \frac{\sin 3x}{3}, \dots, \frac{\sin nx}{n}, \dots$$

(рис. 324), которая сходится к функции $f(x) \equiv 0$ равномерно на любом отрезке $[a, b]$ между тем как последовательность производных этих функций:

$$\cos x, \cos 2x, \dots, \cos nx, \dots$$

расходится в точках $x = (2k+1)\pi$ (k — любое целое число).

Теорема. Если последовательность непрерывно дифференцируемых функций $f_1(x), f_2(x), \dots, f_n(x), \dots$ сходится к функции $f(x)$ на отрезке $[a, b]$, а последовательность производных этих функций

$$f'_1(x), f'_2(x), \dots, f'_n(x), \dots$$

равномерно сходится на этом отрезке, то ее предел равен $f'(x)$.

Итак, дано: $f_n(x) \rightarrow f(x)$ на отрезке $[a, b]$,

$$f'_n(x) \rightarrow F(x) \text{ равномерно на отрезке } [a, b].$$

Докажем, что $F(x) = f'(x)$ или, что то же, докажем, что

$$\lim_{n \rightarrow \infty} f'_n(x) = f'(x) = [\lim_{n \rightarrow \infty} f_n(x)]'.$$

Таким образом, если условия теоремы выполняются, то знаки производной и предела можно переставлять.

Доказательство. Пусть x_0 — какая-нибудь фиксированная точка отрезка $[a, b]$, а x — любая точка этого отрезка. Последовательность $\{f'_n(x)\}$ равномерно сходится к функции $F(x)$ на отрезке с концами x_0 и x , поэтому по теореме об интегрировании равномерно сходящейся последовательности имеем:

$$\int_{x_0}^x f'_n(x) dx \rightarrow \int_{x_0}^x F(x) dx$$

или

$$f_n(x) - f_n(x_0) \rightarrow \int_{x_0}^x F(x) dx. \quad (2)$$

Так как по условию последовательность $\{f_n(x)\}$ сходится к $f(x)$ на отрезке $[a, b]$, то

$$f_n(x) - f_n(x_0) \rightarrow f(x) - f(x_0).$$

В силу единственности предела последовательности, будем иметь:

$$f(x) - f(x_0) = \int_{x_0}^x F(x) dx,$$

откуда на основании теоремы о дифференцировании интеграла с переменным верхним пределом окончательно получаем $f'(x) = F(x)$, что и требовалось доказать.

Нетрудно видеть, что теорема остается справедливой и в случае, если в ее формулировке заменить «отрезок» на «интервал».

Заметим, что условия доказанной теоремы можно несколько ослабить, заменив условие сходимости последовательности $\{f_n(x)\}$ на отрезке $[a, b]$ условием сходимости этой последовательности лишь в какой-нибудь точке рассматриваемого отрезка.

Действительно, если с самого начала проследить доказательство этой теоремы, то нетрудно увидеть, что до соотношения (2) нигде не использовалась сходимос

последовательности $\{f_n(x)\}$. Если за точку x_0 взять данную точку сходимости последовательности $\{f_n(x)\}$ и положить, что в ней предел последовательности равен A , $f_n(x_0) \rightarrow A$, то из соотношения (2) следует:

$$f_n(x) \rightarrow A + \int_{x_0}^x F(x) dx, \quad (3)$$

где x — любая точка отрезка $[a, b]$, а это означает, что последовательность $\{f_n(x)\}$ сходится на отрезке $[a, b]$. Обозначим ее предел через $f(x)$, тогда из соотношения (3) получим $f(x) = A + \int_{x_0}^x F(x) dx$ ($a \leq x \leq b$), откуда $f'(x) = F(x)$ на отрезке $[a, b]$, что и требовалось доказать.

Как следствие доказанной теоремы получаем теорему для ряда. Если члены ряда

$$u_1(x) + u_2(x) + \dots + u_n(x) + \dots$$

— непрерывно дифференцируемые функции на отрезке $[a, b]$ и ряд сходится к функции $f(x)$ на отрезке $[a, b]$, а ряд $u'_1(x) + u'_2(x) + \dots + u'_n(x) + \dots$ равномерно сходится на этом отрезке, то его сумма равна $f'(x)$, т. е.

$$f'(x) = \sum_{n=1}^{\infty} u'_n(x). \quad (4)$$

Заметим, что равенство (4) можно записать следующим образом:

$$\left(\sum_{n=1}^{\infty} u_n(x) \right)' = \sum_{n=1}^{\infty} u'_n(x). \quad (5)$$

Если ряд $\sum_{n=1}^{\infty} u_n(x)$ на отрезке $[a, b]$ удовлетворяет соотношению (5), то говорят, что ряд этот можно *почленно дифференцировать*.

Таким образом, если условия теоремы выполняются, то ряд можно почленно дифференцировать.

Действительно, из условия теоремы получаем: $S_n(x) \rightarrow f(x)$ на отрезке $[a, b]$, и последовательность $\{S'_n(x)\}$ равномерно сходится на этом отрезке.

Так как

$$S'_n(x) = u'_1(x) + u'_2(x) + \dots + u'_n(x)$$

и по доказанной теореме для последовательности

$$S'_n(x) \rightarrow f'(x)$$

на отрезке $[a, b]$, то окончательно получим:

$$\sum_{n=1}^{\infty} u'_n(x) = f'(x),$$

что и требовалось доказать.

Таким образом, если условия теоремы выполняются, то ряд ведет себя аналогично конечной сумме: производная суммы функций равна сумме производных этих функций.

Пример 1. Ряд $\sum_{n=1}^{\infty} \frac{\sin nx}{n^3}$ сходится на всей числовой оси, так как $\left| \frac{\sin nx}{n^3} \right| \leq \frac{1}{n^3}$

для любого значения x . Обозначим сумму этого ряда через $f(x)$. Ряд $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$, полученный из данного дифференцированием его членов, равномерно сходится на всей

числовой оси, так как $\left| \frac{\cos nx}{n^2} \right| \leq \frac{1}{n^2}$, т. е. ряд $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$ мажорируется на промежутке $(-\infty, +\infty)$ числовым сходящимся рядом $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Следовательно, по доказанной теореме:

$$f'(x) = \sum_{n=1}^{\infty} \frac{\cos nx}{n^2}.$$

Пример 2. Рассмотрим пример всюду расходящегося ряда, для которого ряд из производных его членов равномерно сходится на всей числовой оси. Ряд $\sum_{n=1}^{\infty} \cos \frac{x}{n^2}$ удовлетворяет этим условиям.

Действительно, он всюду расходится (так как общий член ряда $\cos \frac{x}{n^2} \rightarrow \cos 0 = 1$), а ряд из производных членов данного ряда $\sum_{n=1}^{\infty} -\frac{\sin \frac{x}{n^2}}{n^2}$ равномерно сходится на всей числовой оси (так как он мажорируется сходящимся рядом $\sum_{n=1}^{\infty} \frac{1}{n^2}$).

Заметим, как и в случае последовательности, что если ряд сходится хотя бы в одной точке x_0 , а ряд из производных его членов равномерно сходится на некотором отрезке, содержащем точку x_0 , то ряд сходится на этом отрезке и можно показать, что сходимость будет равномерной на рассматриваемом отрезке.

§ 8. Понятие о функциональных рядах в комплексной области

Комплексная переменная w называется *функцией комплексной переменной* (аргумента) z [обозначается $w = f(z)$, или $w = \varphi(z)$ и т. д.], если каждому значению z из некоторого множества комплексных чисел по определенному закону ставится в соответствие комплексное число — значение переменной w .

При этом множество значений, принимаемых переменной z , называется *областью определения функции* $w = f(z)$ ¹⁾.

В частности, областью определения функции $w = f(z)$ может быть вся плоскость, полуплоскость, круг, часть плоскости, ограниченная некоторой замкнутой кривой, и др.

Рассмотрим ряд

$$u_1(z) + u_2(z) + \dots + u_n(z) + \dots \quad (1)$$

или, короче, $\sum_{n=1}^{\infty} u_n(z)$, членами которого являются функции комплексной переменной z , определенные в некоторой области D . Возьмем какую-нибудь точку z_0 из области D и подставим ее в ряд (1), тогда получим числовой ряд $\sum_{n=1}^{\infty} u_n(z_0)$. Если ряд $\sum_{n=1}^{\infty} u_n(z_0)$ сходится, то говорят, что ряд (1) сходится в точке z_0 .

Совокупность всех значений z , при которых ряд (1) сходится, называется *областью сходимости* этого ряда.

¹⁾ Если вместо комплексной переменной z рассматривать действительную переменную t , то придем к понятию комплексной функции действительного аргумента t .

Таким образом, для каждой точки z области сходимости ряда (1) определяется число — сумма ряда, которая зависит от рассматриваемой точки z .

Следовательно, в области сходимости ряда (1) определится функция, обозначим ее через $f(z)$, — сумма ряда (1). Таким образом, $f(z) = \sum_{n=1}^{\infty} u_n(z)$ в области сходимости ряда.

Заканчивая на этом рассмотрение произвольных функциональных рядов, переходим к изучению специальных видов функциональных рядов, а именно, степенных и тригонометрических. Эти ряды, как наиболее изученные и наиболее простые в смысле их практического применения, получили широкое распространение.

Раздел 3 СТЕПЕННЫЕ РЯДЫ

§ 9. Интервал сходимости и характер сходимости степенного ряда

Степенным рядом называется ряд вида:

$$a_0 + a_1(z-a) + a_2(z-a)^2 + \dots + a_n(z-a)^n + \dots$$

или, короче,

$$\sum_{n=0}^{\infty} a_n(z-a)^n, \quad (1)$$

где $a_n (n=0, 1, 2, \dots)$ — комплексные числа, называемые *коэффициентами ряда*, a — некоторое комплексное число, z — комплексная переменная. Степенной ряд (1) называют еще рядом по степеням разности $z-a$.

Если $a=0$, то получим степенной ряд вида

$$\sum_{n=0}^{\infty} a_n z^n \quad (2)$$

или ряд по степеням z . Ряд (1) приводится к ряду (2) с помощью подстановки $z-a=z'$. По этой причине мы и будем сначала изучать степенной ряд вида $\sum_{n=0}^{\infty} a_n z^n$.

Если допустить только действительные значения для коэффициентов ряда и переменной, то получим степенной ряд в действительной области,

или *действительный степенной ряд*, он имеет вид $\sum_{n=0}^{\infty} a_n x^n$.

Будем сначала рассматривать комплексный степенной ряд $\sum_{n=0}^{\infty} a_n z^n$. Область сходимости такого ряда устанавливается с помощью следующей теоремы Абеля¹⁾.

Если степенной ряд $\sum_{n=0}^{\infty} a_n z^n$ сходится при $z=z_0$ ($z_0 \neq 0$), то он сходится, и притом абсолютно, для всех z , удовлетворяющих условию $|z| < |z_0|$.

Или, другими словами, из сходимости ряда в точке z_0 следует его абсолютная сходимость внутри круга с центром в начале координат и радиусом $|z_0|$ (рис. 437).

¹⁾ Абель (1802—1829) — норвежский математик.

Доказательство. Так как ряд $\sum_{n=0}^{\infty} a_n z_0^n$ сходится, то $a_n z_0^n \rightarrow 0$ и, следовательно, найдется такое число $M > 0$, что при всех n будет выполняться неравенство $|a_n z_0^n| < M$ (члены ряда ограничены). Рассмотрим тождество $|a_n z^n| = |a_n z_0^n| \left| \frac{z}{z_0} \right|^n$. Так как оно выполняется при любом z , то при любом z имеем:

$$|a_n z^n| < M \left| \frac{z}{z_0} \right|^n.$$

Пусть z — любое число, для которого $|z| < |z_0|$, тогда $\left| \frac{z}{z_0} \right| = q$, $0 < q < 1$, и, следовательно, $|a_n z^n| < M q^n$. Так как ряд $\sum_{n=0}^{\infty} M q^n$ ($0 < q < 1$) сходится, то по признаку сравнения рядов сходится и ряд

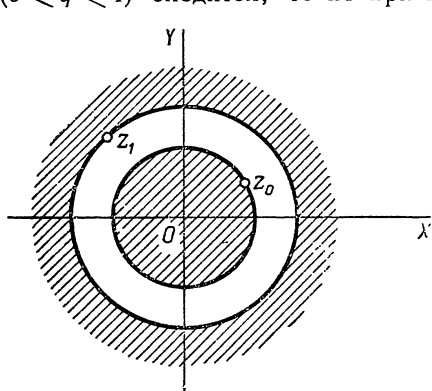


Рис. 437.

$\sum_{n=0}^{\infty} |a_n z^n|$ или, что то же, ряд $\sum_{n=0}^{\infty} a_n z^n$ абсолютно сходится.

Но z выбиралось произвольно, лишь бы удовлетворялось условие $|z| < |z_0|$. Учитывая это, заключаем, что степенной ряд $\sum_{n=0}^{\infty} a_n z^n$ абсолютно сходится в любой точке z , удовлетворяющей условию $|z| < |z_0|$, что и требовалось доказать.

Следствие. Если степенной ряд расходится или неабсолютно сходится при $z = z_1$, то он расходится и при всех z , удовлетворяющих условию $|z| > |z_1|$. Или, что то же, из расходимости (или неабсолютной сходимости) степенного ряда в точке z_1 следует расходимость его вне круга радиуса $|z_1|$ с центром в начале координат (рис. 437).

В самом деле, допустив сходимость ряда в какой-нибудь точке z_2 , для которой $|z_2| > |z_1|$, получим по теореме Абеля абсолютную сходимость ряда во всех точках z , для которых $|z| < |z_2|$, в частности и в точке z_1 , что противоречит условию. Следовательно, в точке z_2 степенной ряд расходится. Полученное противоречие доказывает следствие.

Ниже будет рассмотрен вопрос об области сходимости степенного ряда $\sum_{n=0}^{\infty} a_n z^n$.

Прежде всего заметим, что в точке $z = 0$ каждый степенной ряд сходится и его сумма равна a_0 . При этом возможны следующие случаи.

I. Степенной ряд $\sum_{n=0}^{\infty} a_n z^n$ может сходиться только в одной точке $z = 0$.

Например, ряд $\sum_{n=1}^{\infty} n^n z^n$ сходится только при $z = 0$, так как если $z \neq 0$, то $|n^n z^n| = (n|z|)^n$ не стремится к нулю ($n|z| > 1$, начиная с некоторого n).

II. Степенной ряд может сходиться при любом значении z , т. е. во всей плоскости.

Например, возьмем ряд $1 + \frac{z}{1!} + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots$. Так как

$$\lim_{n \rightarrow \infty} \frac{|z|^{n+1} n!}{(n+1)! |z|^n} = |z| \lim_{n \rightarrow \infty} \frac{1}{(n+1)} = 0 < 1,$$

то по признаку Даламбера ряд сходится при любом z .

III. Степенной ряд может иметь точку сходимости, отличную от нуля, и точку расходимости. Например, ряд геометрической прогрессии $1 + z + z^2 + \dots + z^n + \dots$ сходится при $z = \frac{1}{2}$ и расходится при $z = 2$.

Последний случай рассмотрим более подробно. Пусть степенной ряд $\sum_{n=0}^{\infty} a_n z^n$ имеет точку сходимости, отличную от нуля, и точку расходимости, следовательно, по теореме Абеля и следствию из нее существует бесчисленное множество точек сходимости и точек расходимости, в частности, существуют положительные точки сходимости и точки расходимости.

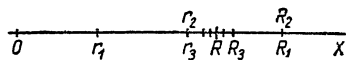


Рис. 438.

Возьмем какую-нибудь положительную точку сходимости r_1 и положительную точку расходимости R_1 . Ясно, что $r_1 < R_1$

(рис. 438). Если $\frac{r_1 + R_1}{2}$ является точкой сходимости, то положим $r_2 = \frac{r_1 + R_1}{2}$, $R_2 = R_1$, если же $\frac{r_1 + R_1}{2}$ — точка расходимости, то положим $r_2 = r_1$, $R_2 = \frac{r_1 + R_1}{2}$. Аналогично введем числа r_3 , R_3 и т. д., получим две монотонные последовательности:

$$r_1 \leq r_2 \leq r_3 \leq \dots$$

и

$$R_1 \geq R_2 \geq R_3 \geq \dots$$

Первая последовательность является последовательностью точек сходимости ряда, вторая — последовательностью точек расходимости ряда. Эти последовательности сходятся, как монотонные и ограниченные, при этом они имеют общий предел, так как $R_n - r_n = \frac{R_1 - r_1}{2^{n-1}} \rightarrow 0$.

Обозначим через $R = \lim r_n = \lim R_n$ и докажем, что в круге $|z| < R$ ряд сходится, и притом абсолютно, вне этого круга, т. е. для $|z| > R$, ряд расходится.

Действительно, пусть точка z лежит внутри круга $|z| < R$, тогда при достаточно большом n выполняется условие $|z| < r_n$. Если учесть, что r_n — точка сходимости ряда, то по теореме Абеля ряд абсолютно сходится в точке z .

Если же точка z лежит вне круга, т. е. $|z| > R$, то при достаточно большом n имеет место неравенство $|z| > R_n$. Так как в точке R_n ряд расходится, то по следствию из теоремы Абеля ряд расходится и в точке z .

Таким образом, ряд сходится абсолютно внутри круга радиуса R с центром в точке $z = 0$ и расходится вне его. Этот круг называется *кругом сходимости степенного ряда*. Говоря о круге сходимости, будем иметь в виду только внутренние точки этого круга. Радиус круга сходимости называется *радиусом сходимости степенного ряда*. О сходимости степенного ряда на окружности $|z| = R$ (R — радиус сходимости) в общем случае ничего сказать нельзя. Здесь, в зависимости от ряда, могут представиться различные случаи.

Таким образом, областью сходимости степенного ряда $\sum_{n=0}^{\infty} a_n z^n$ является круг радиуса R с центром в точке $z=0$ (точнее, внутренность этого круга и, быть может, некоторое множество точек окружности $|z|=R$).

Если степенной ряд $\sum_{n=0}^{\infty} a_n z^n$, как в примере 1, сходится только в одной точке ($z=0$), то полагают $R=0$, если же степенной ряд сходится в любой точке z , то полагают $R=\infty$ и говорят, что круг сходимости степенного ряда в первом случае вырождается в точку, а во втором случае является всей плоскостью.

Для практического определения радиуса сходимости степенного ряда в простейших случаях применяют признак Даламбера или Коши к ряду из модулей его членов. Последнее возможно, так как степенной ряд абсолютно сходится внутри круга сходимости и расходится вне его.

Пример. Найти радиус сходимости ряда $\sum_{n=1}^{\infty} \left(2i + \frac{1}{n}\right)^n z^{2n}$.

Решение. Применяя признак сходимости Коши, имеем:

$$\lim_{n \rightarrow \infty} \sqrt[n]{\left|2i + \frac{1}{n}\right|^n |z|^{2n}} = |z|^2 \lim_{n \rightarrow \infty} \left|2i + \frac{1}{n}\right| = |z|^2 \lim_{n \rightarrow \infty} \sqrt{\frac{1}{n^2} + 4} = 2|z|^2.$$

Если $2|z|^2 < 1$, то ряд сходится. Если $2|z|^2 > 1$, то ряд расходится, следовательно, при $|z| < \frac{\sqrt{2}}{2}$ ряд сходится, а при $|z| > \frac{\sqrt{2}}{2}$ ряд расходится, т. е. $R = \frac{\sqrt{2}}{2}$.

Если степенной ряд $\sum_{n=0}^{\infty} a_n z^n$ сходится в круге радиуса $R (R > 0)$, то его сумма является функцией, обозначим ее через $f(z)$, *комплексной переменной* z :

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad |z| < R.$$

Если же z будет принимать только действительные значения $z=x$, то ряд $\sum_{n=0}^{\infty} a_n x^n$ будет сходиться в общей части круга сходимости и действительной оси, т. е. на интервале $(-R, R)$ (о сходимости ряда на концах интервала ничего определенного сказать нельзя.) В этом случае сумма ряда является функцией *действительной переменной* x , которая может быть и комплекснозначной, т. е. может принимать комплексные значения, если не все коэффициенты ряда — действительные числа.

Если же и коэффициенты ряда и переменная z действительны, то ряд, как было уже замечено, называется действительным и его суммой является действительная функция действительной переменной, определенная в области сходимости ряда.

Для нахождения области сходимости действительного ряда $\sum_{n=0}^{\infty} a_n x^n$, кроме определения *интервала сходимости*, нужно еще исследовать сходимость ряда на концах этого интервала, т. е. подставить в ряд значения $x=-R$, $x=R$ и исследовать сходимость соответствующих числовых рядов.

Далее в этом разделе будут рассматриваться только действительные степенные ряды.

Пример. Найти область сходимости ряда $1 + \frac{x}{1} + \frac{x^2}{2} + \dots + \frac{x^n}{n} + \dots$

Решение. Применяя признак Даламбера, получим: $\lim_{n \rightarrow \infty} \frac{|x|^{n+1}n}{(n+1)|x|^n} = |x|$, следовательно, ряд сходится при $|x| < 1$ и расходится при $|x| > 1$.

Таким образом, радиус сходимости ряда $R=1$ и, следовательно, интервал сходимости $(-1, 1)$. Исследуем поведение ряда на концах интервала сходимости. Подставляя в ряд значения $x=1$ и $x=-1$, получим числовые ряды

$$1 + 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots,$$

$$1 - 1 + \frac{1}{2} - \frac{1}{3} + \dots + \frac{(-1)^n}{n} + \dots,$$

первый из которых — гармонический ряд — расходится, второй — сходится по признаку Лейбница. Следовательно, область сходимости является полуинтервалом $[-1, 1)$.

Характер сходимости степенного ряда

Теорема. Степенной ряд сходится равномерно на любом отрезке, лежащем внутри интервала сходимости.

Действительно, пусть отрезок $[a, b]$ лежит внутри интервала сходимости $(-R, R)$ ряда $\sum_{n=0}^{\infty} a_n x^n$. Тогда найдется такое число r , $0 < r < R$,

что отрезок $[-r, r]$ будет содержать отрезок $[a, b]$. Ряд $\sum_{n=0}^{\infty} a_n x^n$ абсо-

лютно сходится при $x=r$, т. е. сходится ряд $\sum_{n=0}^{\infty} |a_n| r^n$. Но для всех x отрезка $[-r, r]$ выполняется неравенство $|a_n x^n| \leq |a_n| r^n$, которое означает, что ряд $\sum_{n=0}^{\infty} a_n x^n$ на отрезке $[-r, r]$ мажорируется числовым сходящимся рядом.

Следовательно, по признаку Вейерштрасса ряд $\sum_{n=0}^{\infty} a_n x^n$ равномерно сходится на этом отрезке, в частности, он равномерно сходится и на отрезке $[a, b]$, что и требовалось доказать.

Заметим, что сходимость ряда на его интервале сходимости $(-R, R)$ может быть и неравномерной. Так например, ряд $\sum_{n=0}^{\infty} x^n$ на интервале сходимости $(-1, 1)$ сходится неравномерно.

Как следствие доказанной теоремы и теоремы о непрерывности суммы равномерно сходящегося ряда (§ 6) получаем **теорему:** сумма степенного ряда является непрерывной функцией на интервале сходимости.

Действительно, пусть x_0 — произвольная точка интервала сходимости ряда $\sum_{n=0}^{\infty} a_n x^n$. Рассмотрим какой-нибудь отрезок $[a, b]$, принадлежащий интервалу сходимости, внутри которого находится точка x_0 .

Ряд $\sum_{n=0}^{\infty} a_n x^n$ равномерно сходится на отрезке $[a, b]$, члены ряда $a_n x^n$ ($n=0, 1, 2, \dots$) — непрерывные функции на этом отрезке, следовательно, сумма ряда — непрерывная функция на отрезке $[a, b]$. В частности, сумма ряда — непрерывная функция в точке x_0 . В силу произвольности выбора точки x_0 из интервала $(-R, R)$ получаем, что сумма ряда — непрерывная функция на интервале сходимости.

§ 10. Интегрирование и дифференцирование степенных рядов

Теорема. Степенной ряд можно почленно интегрировать по любому отрезку $[a, b]$, принадлежащему интервалу сходимости ряда.

В частности, степенной ряд можно почленно интегрировать в пределах от 0 до x , где $|x| < R$ (R — радиус сходимости ряда).

В самом деле, степенной ряд равномерно сходится на отрезке $[a, b]$, принадлежащем интервалу сходимости ряда. Заключение теоремы следует немедленно, если принять во внимание теорему (см. § 7) об интегрировании равномерно сходящихся рядов.

В качестве примеров на интегрирование степенных рядов рассмотрим разложение в степенные ряды функций $\operatorname{arctg} x$ и $\ln(1+x)$.

Действительно,

$$\operatorname{arctg} x = \int_0^x \frac{dx}{1+x^2}.$$

Как известно,

$$\frac{1}{1-x} = 1 + x + x^2 + \dots + x^n + \dots \quad (1)$$

при $|x| < 1$. Подставляя в это равенство $-x^2$ вместо x , получим равенство $\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots + (-1)^n x^{2n} + \dots$, справедливое при $|-x^2| < 1$ или, что то же, при $|x| < 1$. Если почленно проинтегрировать этот ряд в пределах от 0 до x , $|x| < 1$, то получим:

$$\int_0^x \frac{dx}{1+x^2} = \operatorname{arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots + (-1)^{n+1} \frac{x^{2n-1}}{(2n-1)} + \dots,$$

откуда

$$\operatorname{arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots + (-1)^{n+1} \frac{x^{2n-1}}{2n-1} + \dots \quad (|x| < 1). \quad (2)$$

Аналогично получим разложение в степенной ряд функции $\ln(1+x)$.

Так как $\ln(1+x) = \int_0^x \frac{dx}{1+x}$, то, подставляя $-x$ вместо x в равенство (1), получим:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-1)^n x^n + \dots \quad (|x| < 1).$$

Интегрируя полученное равенство в пределах от 0 до x , $|x| < 1$, запишем:

$$\int_0^x \frac{dx}{1+x} = \int_0^x (1 - x + x^2 - x^3 + \dots) dx$$

или

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^{n+1} \frac{x^n}{n} + \dots \quad (3)$$

при $|x| < 1$.

Ряд (3) является знакочередующимся для x из интервала $(0, 1)$. А, как известно, для знакочередующегося ряда, удовлетворяющего условиям признака Лейбница, ошибка при замене суммы ряда его частной суммой по абсолютной величине не больше абсолютной величины первого из отброшенных членов ряда, т. е. в этом случае ошибку можно оценить.

Но, несмотря на это, оказывается, что ряд (3) неудобен для вычисления логарифмов чисел, так как его члены медленно убывают. Для

вычисления логарифмов чисел применяется другой ряд, который мы и получим. Подставляя в ряд (3) $-x$ вместо x , имеем:

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots - \frac{x^n}{n} - \dots \quad (|x| < 1). \quad (4)$$

Вычитая из ряда (3) ряд (4), получим:

$$\ln\left(\frac{1+x}{1-x}\right) = 2x + 2\frac{x^3}{3} + 2\frac{x^5}{5} + \dots + 2\frac{x^{2n-1}}{2n-1} + \dots \quad (|x| < 1).$$

Если теперь положить $x = \frac{1}{2N+1}$ (где N — натуральное число, следовательно, $0 < x < 1$), то

$$\frac{1+x}{1-x} = \frac{1 + \frac{1}{2N+1}}{1 - \frac{1}{2N+1}} = \frac{N+1}{N},$$

и тогда $\ln \frac{N+1}{N} = \frac{2}{2N+1} + \frac{2}{3(2N+1)^3} + \frac{2}{5(2N+1)^5} + \dots$, откуда

$$\ln(N+1) = \ln N + \frac{2}{2N+1} + \frac{2}{3(2N+1)^3} + \frac{2}{5(2N+1)^5} + \dots \quad (5)$$

Равенство (5) дает возможность последовательно находить логарифмы натуральных чисел. Не останавливаясь на этом, заметим только, что если заменить сумму ряда (5) какой-нибудь его частной суммой, то легко оценить величину остатка ряда. (Остаток ряда в этом случае равен сумме ряда с членами, меньшими членов ряда геометрической прогрессии со знаменателем $q = \frac{1}{(2N+1)^2}$, $0 < q < 1$, а известно, что сумма ряда $q^k + q^{k+1} + \dots$ равна $\frac{q^k}{1-q}$.)

Рассмотрим теперь дифференцирование степенных рядов.

Пусть степенной ряд

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots \quad (6)$$

сходится в интервале $(-R, R)$, а ряд

$$a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1} + \dots \quad (7)$$

получен из ряда (6) дифференцированием его членов.

Лемма. Ряд (7) имеет тот же интервал сходимости, что и ряд (6).

Доказательство. Пусть x_0 произвольная точка интервала сходимости ряда (6), т. е. $|x_0| < R$. Тогда возьмем число r , удовлетворяющее условию $|x_0| < r < R$. Оценим модуль общего члена ряда (7) в рассматриваемой точке x_0 .

$$\text{Нетрудно видеть, что } |na_nx_0^{n-1}| = |a_nr^n| \cdot \frac{n|x_0|^{n-1}}{r^n}.$$

Ряд $\sum_{n=0}^{\infty} a_nx^n$ абсолютно сходится в точке r , т. е. ряд $\sum_{n=0}^{\infty} |a_n|r^n$ сходится, и, следовательно, общий член ряда $|a_n|r^n$ стремится к нулю. Это значит, что члены этого ряда ограничены, $|a_n|r^n < M$, где M — некоторое число.

Учитывая это, получим:

$$|na_nx_0^{n-1}| < M \frac{n|x_0|^{n-1}}{r^n}.$$

Ряд $\sum_{n=1}^{\infty} M \frac{n|x_0|^{n-1}}{r^n}$ сходится (на основании признака Даламбера), так как

$$\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \frac{(n+1)|x_0|}{nr} = \frac{|x_0|}{r} < 1.$$

Используя признак сравнения рядов, получаем сходимость ряда $\sum_{n=1}^{\infty} |na_n x_0^{n-1}|$ или, что то же, абсолютную сходимость ряда $\sum_{n=1}^{\infty} na_n x^{n-1}$ в точке x_0 . В силу произвольности выбора точки x_0 из интервала $(-R, R)$ заключаем, что ряд (7) имеет интервал сходимости, не меньший интервала сходимости ряда (6).

Докажем теперь, что интервал сходимости ряда (7) не больше интервала сходимости ряда (6).

Допустим противное, а именно, допустим, что ряд (7) сходится в интервале $(-R_1, R_1)$, $R_1 > R$. Тогда, так как ряд (7) степенной, то его можно почленно интегрировать в пределах от 0 до x , где $|x| < R_1$. Ряд, полученный в результате интегрирования, сходится в интервале $(-R_1, R_1)$, он отличается от ряда (6) только постоянным слагаемым. Следовательно, и ряд (6) сходится в интервале $(-R_1, R_1)$, что противоречит допущению.

Таким образом, отвергая допущение, получим утверждение леммы.

Теорема. Степенной ряд в интервале сходимости можно почленно дифференцировать, т. е. если $f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$ в интервале $(-R, R)$, то $f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \dots + na_n x^{n-1} + \dots$ в том же интервале.

В самом деле, пусть x_0 — произвольная точка интервала $(-R, R)$. Возьмем какой-нибудь отрезок $[a, b]$, внутри которого находится точка x_0 , принадлежащий интервалу $(-R, R)$.

Ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится к $f(x)$ на отрезке $[a, b]$, ряд из производных его членов равномерно сходится на этом отрезке [так как это степенной ряд с интервалом сходимости $(-R, R)$]. Тогда, по теореме дифференцирования функциональных рядов, сумма ряда из производных равна $f'(x)$ на отрезке $[a, b]$, а следовательно, и в точке x_0 . Но так как точка x_0 выбиралась произвольно из интервала $(-R, R)$, то теорема доказана.

Применяя последовательно эту теорему несколько раз, получим:

Следствие 1. Степенной ряд в интервале сходимости можно почленно дифференцировать любое число раз, т. е. если $f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$ в интервале $(-R, R)$, то

$$\begin{aligned} f'(x) &= a_1 + 2a_2 x + 3a_3 x^2 + \dots + na_n x^{n-1} + \dots \\ f''(x) &= 2a_2 + 2 \cdot 3a_3 x + \dots + (n-1)na_n x^{n-2} + \dots \\ &\dots \dots \dots \end{aligned}$$

причем ряды, полученные в результате дифференцирования, сходятся в том же интервале $(-R, R)$.

Следствие 2. Сумма степенного ряда — бесконечно дифференцируемая функция в интервале сходимости ряда.

Функция $f(x)$ называется *бесконечно дифференцируемой в точке (на интервале)*, если она имеет производные любого порядка в этой точке (на этом интервале).

Рассмотрим теперь степенной ряд по степеням разности $z - a$.

$$a_0 + a_1(z - a) + a_2(z - a)^2 + \dots + a_n(z - a)^n + \dots \quad (8)$$

Как было указано, этот ряд с помощью подстановки $z - a = z'$ приводится к виду:

$$a_0 + a_1 z' + a_2 (z')^2 + \dots + a_n (z')^n + \dots \quad (9)$$

Если же коэффициенты ряда (8), число a и переменная z действительные, то ряд (8) примет вид:

$$\sum_{n=0}^{\infty} a_n (x-a)^n. \quad (8')$$

И тогда подстановка $x - a = x'$ преобразует этот ряд к виду

$$\sum_{n=0}^{\infty} a_n (x')^n. \quad (9')$$

При этом, если последний ряд сходится в интервале $(-R, R)$ (R —

Рис. 439.

радиус сходимости ряда), то ряд $\sum_{n=0}^{\infty} a_n(x-a)^n$ будет сходиться в интервале $(a-R, a+R)$, симметричном относительно точки a (рис. 439).

Учитывая это, заметим, что все сказанное о степенных рядах по степеням x остается справедливым и для степенных рядов по степеням разности $x - a$ в соответствующих интервалах сходимости.

§ 11. Ряд Тейлора. Основные разложения

Пусть функция $f(x)$ является суммой степенного ряда по степеням разности $x - a$, сходящегося в некотором интервале:

$$f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \dots + a_n(x-a)^n + \dots$$

Тогда в интервале сходимости, который, как известно, симметричен относительно точки a , ряд можно почленно дифференцировать любое число раз.

Дифференцируя, получим:

$$\begin{aligned} f'(x) &= 1 \cdot a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + 4a_4(x-a)^3 + \dots + \\ &\quad + na_n(x-a)^{n-1} + \dots \\ f''(x) &= 1 \cdot 2a_2 + 2 \cdot 3a_3(x-a) + 3 \cdot 4a_4(x-a)^2 + 4 \cdot 5a_5(x-a)^3 + \dots + \\ &\quad + n(n-1)a_n(x-a)^{n-2} + \dots \\ f'''(x) &= 1 \cdot 2 \cdot 3a_3 + 2 \cdot 3 \cdot 4a_4(x-a) + \dots + \\ &\quad + n(n-1)(n-2)a_n(x-a)^{n-3} + \dots \\ &\quad \vdots \\ f^{(n)}(x) &= 1 \cdot 2 \cdot 3 \dots na_n + 2 \cdot 3 \cdot 4 \dots (n+1)a_{n+1}(x-a) + \dots \end{aligned}$$

Полагая в этих равенствах $x=a$, получим:

$$f(a) = a_0, f'(a) = 1 \cdot a_1, f''(a) = 2!a_2, f'''(a) = 3!a_3, \dots, f^{(n)}(a) = n!a_n, \dots,$$

откуда находим:

$$a_0 = f(a), \quad a_1 = \frac{f'(a)}{1!}, \quad a_2 = \frac{f''(a)}{2!},$$

$$a_3 = \frac{f'''(a)}{3!}, \quad \dots, \quad a_n = \frac{f^{(n)}(a)}{n!}, \quad \dots$$

Подставляя полученные выражения коэффициентов в ряд для $f(x)$, запишем:

$$f(x) = f(a) + \frac{f'(a)}{1!} (x-a) + \frac{f''(a)}{2!} (x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!} (x-a)^n + \dots$$

Рядом Тейлора функции $f(x)$ по степеням $x - a$ или, что то же, в окрестности точки a называется степенной ряд вида:

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \dots$$

(при этом предполагается, что функция $f(x)$ бесконечно дифференцируема в точке a).

При $a=0$ получим ряд по степеням x , который называется *рядом Маклорена функции $f(x)$* .

Приведенными выше рассуждениями доказана **теорема**: если функция $f(x)$ в некоторой окрестности точки a является суммой степенного ряда по степеням $x - a$, то этот ряд является рядом Тейлора функции $f(x)$.

Из этой теоремы как следствие получается **теорема единственности**: если функция $f(x)$ в некоторой окрестности точки a разложена в степенной ряд по степеням $x - a$, т. е. функция $f(x)$ является суммой сходящегося в некоторой окрестности a ряда по степеням $x - a$, то такое разложение единственное. Или иначе: не может быть двух различных степенных рядов по степеням $x - a$, сходящихся к одной и той же функции.

В самом деле, это получается сразу, если учесть, что коэффициенты ряда выражаются через значения функции $f(x)$ и ее производных в точке a .

Заметим, что при практическом применении этой теоремы, в частности при решении дифференциальных уравнений с помощью степенных рядов, используют следующую формулировку этой **теоремы**: если имеются два тождественно равных сходящихся на некотором интервале ряда¹⁾ по степеням $x - a$, то коэффициенты этих рядов при одинаковых степенях $x - a$ равны.

Таким образом, теорема единственности дает обоснование метода сравнения коэффициентов, который неоднократно использовался в курсе, в частности при разложении рациональной дроби на сумму элементарных дробей и при решении дифференциальных уравнений методом подбора.

До сих пор мы определяли область сходимости степенного ряда и выясняли свойства суммы ряда в его интервале сходимости. Будем решать теперь следующую задачу: дана функция и требуется разложить ее в степенной ряд. Отчасти такую задачу мы уже решали, когда с помощью интегрирования степенных рядов разлагали в ряд по степеням x функции $\arctg x$ и $\ln(1+x)$. Как известно, если функция $f(x)$ является суммой степенного ряда, то в интервале сходимости ряда функция $f(x)$ имеет производные любого порядка. Отсюда следует, что ставить задачу о разложении функции $f(x)$ в ряд по степеням $x - a$ можно лишь по отношению к функции, имеющей производные любого порядка в некоторой окрестности точки a . Если еще учесть, что степенной ряд, сходящийся к функции $f(x)$, является рядом Тейлора этой функции, то поставленная задача сводится к разложению функции $f(x)$ в окрестности точки a в ряд Тейлора.

Пусть функция $f(x)$ в некоторой окрестности точки a имеет производные любого порядка. Тогда для функции $f(x)$ в окрестности a можно записать соответствующий ей ряд Тейлора:

$$f(x) \sim f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \dots$$

($\sim$ — знак соответствия). Вместо знака соответствия можно записать знак $=$ (знак равенства) в окрестности точки a лишь при условии, что ряд Тейлора функции $f(x)$ сходится к этой функции в рассматриваемой окрестности точки a .

¹⁾ Это означает, что суммы этих рядов тождественно равны в интервале сходимости.

Заметим, что ряд Тейлора функции $f(x)$ может и не сходиться к этой функции в окрестности рассматриваемой точки.

Например, функция $f(x)$, определяемая условиями:

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0, \end{cases}$$

бесконечно дифференцируема на всей числовой оси, при этом все ее производные в точке $x=0$ равны нулю.

Действительно, если $x \neq 0$, то $f'(x) = \frac{2}{x^3} e^{-\frac{1}{x^2}}$, дифференцируя эту функцию, найдем $f''(x)$ при $x \neq 0$ и т. д.

Если же $x=0$, то

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(\Delta x) - f(0)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{e^{-\frac{1}{(\Delta x)^2}}}{\Delta x} = 0,$$

$$f''(0) = \lim_{\Delta x \rightarrow 0} \frac{f'(\Delta x) - f'(0)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\frac{2}{(\Delta x)^3} e^{-\frac{1}{(\Delta x)^2}}}{\Delta x} = 0 \text{ и т. д.,}$$

получим, что значения всех производных этой функции при $x=0$ равны нулю.

Следовательно, функции $f(x)$ соответствует ряд по степеням x , все коэффициенты которого равны нулю:

$$f(x) \sim 0 + 0 \cdot x + \dots + 0 \cdot x^n + \dots$$

Его сумма при всех x тождественно равна нулю и, следовательно, равна функции $f(x)$ только в одной точке, при $x=0$.

Разность между функцией $f(x)$ и частной суммой $S_n(x)$ ряда Тейлора функции $f(x)$ называется *остаточным членом ряда Тейлора* и обозначается $R_n(x)$:

$$R_n(x) = f(x) - S_n(x).$$

Как было показано (гл. IV, § 15), функция $R_n(x)$ имеет следующий вид:

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1},$$

где ξ лежит между a и x или, что то же,

$$\xi = a + \theta(x-a), \quad 0 < \theta < 1.$$

Если воспользоваться определениями сходимости и суммы ряда, то получим следующую теорему: для того чтобы ряд Тейлора функции $f(x)$ сходил к ней в некотором интервале, *необходимо и достаточно*, чтобы в рассматриваемом интервале стремился к нулю остаточный член этого ряда.

Таким образом, задача разложения функции $f(x)$ в степенной ряд сведена по существу к определению значений x , при которых $R_n(x) \rightarrow 0$.

Основные разложения. Рассмотрим разложение функций e^x , $\sin x$, $\cos x$ в ряды по степеням x . Для этого запишем ряд Маклорена, соответствующий функции $f(x)$:

$$f(x) \sim f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 + \dots + \frac{f^{(n)}(0)}{n!} x^n + \dots$$

и исследуем, при каких значениях x остаточный член этого ряда

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1},$$

где точка ξ лежит между 0 и x , стремится к нулю.

Действительно, пусть $f(x) = e^x$. Тогда $f'(x) = f''(x) = \dots = f^{(n)}(x) = \dots = e^x$, откуда при $x = 0$ получаем: $f(0) = f'(0) = \dots = f^{(n)}(0) = \dots = 1$, и, следовательно,

$$e^x \sim 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots \quad (1)$$

При этом остаточный член $R_n(x)$ ряда (1) имеет вид:

$$R_n(x) = \frac{e^\xi}{(n+1)!} x^{n+1}, \text{ где } \xi = \theta x, 0 < \theta < 1.$$

Доказав, что $R_n(x) \rightarrow 0$ при любом значении x , мы тем самым докажем, что в соотношении (1) вместо знака соответствия можно при любом x , $-\infty < x < +\infty$, записать знак равенства.

Для этого сначала докажем лемму: для любого x , $-\infty < x < +\infty$, справедливо следующее равенство:

$$\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0.$$

Доказательство. Применим признак Даламбера к ряду $1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$. Так как

$$\lim_{n \rightarrow \infty} \frac{|x|^{n+1} n!}{(n+1)! |x|^n} = |x| \lim_{n \rightarrow \infty} \frac{1}{(n+1)} = |x| \cdot 0 = 0 < 1,$$

то рассматриваемый ряд сходится при всех значениях x , $-\infty < x < +\infty$, и, следовательно, при любом x общий член ряда $\frac{x^n}{n!} \rightarrow 0$, что и требовалось доказать.

Учитывая, что $e^\xi < e^{|x|}$, на основании леммы заключаем, что при любом значении x $R_n(x) = \frac{e^\xi}{(n+1)!} x^{n+1} \rightarrow 0$, поэтому при любом x

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

Пусть теперь $f(x) = \sin x$. Тогда

$$\begin{aligned} f'(x) &= \cos x, \quad f''(x) = -\sin x, \quad f'''(x) = -\cos x, \quad f^{IV}(x) = \sin x, \\ &\dots, \quad f^{(n)}(x) = \sin\left(x + n \frac{\pi}{2}\right), \dots, \end{aligned}$$

откуда, полагая $x = 0$, найдем:

$$\begin{aligned} f(0) &= 0, \quad f'(0) = 1, \quad f''(0) = 0, \quad f'''(0) = -1, \quad f^{IV}(0) = 0, \dots \\ &\dots \end{aligned}$$

Следовательно,

$$\sin x \sim x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^{n+1} \frac{x^{2n-1}}{(2n-1)!} + \dots,$$

при этом

$$R_n(x) = \frac{\sin\left[\xi + (n+1) \frac{\pi}{2}\right]}{(n+1)!} x^{n+1}, \text{ где } \xi = \theta x, 0 < \theta < 1.$$

Но так как $|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!} \rightarrow 0$ при любом x , как это следует из леммы, то при любом x :

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^{n+1} \frac{x^{2n-1}}{(2n-1)!} + \dots$$

Разложение $\cos x$ в степенной ряд можно получить, повторяя предыдущие рассуждения. Можно также получить разложение для $\cos x$ путем дифференцирования ряда для $\sin x$.

Действительно, дифференцируя ряд для $\sin x$, получим при любом x

$$\cos x = 1 - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots$$

Гиперболические функции. Функции $\frac{e^x - e^{-x}}{2}$ и $\frac{e^x + e^{-x}}{2}$ называются соответственно *гиперболическим синусом* и *гиперболическим косинусом* и обозначаются $\operatorname{sh} x$ и $\operatorname{ch} x$ (читаются соответственно: «синус гиперболический x », «косинус гиперболический x »).

Итак, по определению

$$\operatorname{sh} x = \frac{e^x - e^{-x}}{2} \quad \operatorname{ch} x = \frac{e^x + e^{-x}}{2}.$$

Тангенс гиперболический, обозначается $\operatorname{th} x$ (читается: «тангенс гиперболический x »), определяется равенством

$$\operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

Из определения рассматриваемых функций следует, что они существуют при всех значениях x , кроме того, $\operatorname{sh} x$, $\operatorname{th} x$ — нечетные функции, а $\operatorname{ch} x$ — четная функция.

Нетрудно получить основное соотношение:

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1.$$

Не останавливаясь на этом, заметим только, что название «гиперболические функции» указывает на связь этих функций с гиперболой, так же как название «круговые функции» указывает на их связь с окружностью.

Рекомендуется построить графики этих функций, при этом для построения графиков функций $\operatorname{sh} x$ и $\operatorname{ch} x$ можно использовать сложение графиков функций e^x и e^{-x} , а при построении графика функции $\operatorname{th} x$ — использовать ее нечетность, а также следующие ее свойства: 1) $\operatorname{th} 0 = 0$, 2) $0 < \operatorname{th} x < 1$ при $x > 0$ и 3) $\operatorname{th} x \rightarrow 1$ при $x \rightarrow +\infty$.

Из определения функций $\operatorname{sh} x$ и $\operatorname{ch} x$ легко получить их разложения в степенные ряды.

Действительно, так как

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots \quad (1')$$

при любом x , то, подставляя $-x$ вместо x , получим:

$$e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \dots + (-1)^n \frac{x^n}{n!} + \dots \quad (1'')$$

при любом x .

Вычитая и складывая ряды (1') и (1''), получим:

$$\begin{aligned} \operatorname{sh} x &= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + \frac{x^{2n-1}}{(2n-1)!} + \dots \\ \operatorname{ch} x &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots + \frac{x^{2n}}{(2n)!} + \dots \end{aligned}$$

Суммируя полученное, мы можем записать следующие *пять основных разложений*, справедливых при всех значениях x :

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!},$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!},$$

$$\operatorname{sh} x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \dots = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!},$$

$$\operatorname{ch} x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \dots = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}.$$

В заключение заметим, что для практических приложений не всегда удобно производить разложение функций в степенные ряды непосредственно по формуле Тейлора. Это следует, во-первых, из того, что выражение производной любого порядка функции $f(x)$ не всегда можно записать, а во-вторых, даже в том случае, когда такое выражение получено, исследование остаточного члена ряда, как правило, представляет большие трудности. Обычно при разложении функций в степенные ряды стараются обойтись без формулы Тейлора, используя известные разложения²⁾ и операции над степенными рядами. В частности, мы рассматривали интегрирование и дифференцирование степенных рядов, помимо этого на примерах гиперболических функций рассматривалось сложение и вычитание степенных рядов.

§ 12. Операции над степенными рядами

1) Сложение, вычитание и умножение рядов. Даны два ряда

$$\sum_{n=0}^{\infty} a_n x^n \quad (1)$$

и

$$\sum_{n=0}^{\infty} b_n x^n. \quad (2)$$

Пусть они сходятся соответственно в интервалах $(-R, R)$ и $(-R_1, R_1)$ и их суммы соответственно равны $f(x)$ и $\varphi(x)$.

Суммой, разностью, произведением степенных рядов (1) и (2) называются соответственно ряды:

$$\sum_{n=0}^{\infty} (a_n \pm b_n) x^n, \quad (3)$$

$$a_0 b_0 + (a_0 b_1 + a_1 b_0) x + (a_0 b_2 + a_1 b_1 + a_2 b_0) x^2 + \dots + (a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0) x^n + \dots \quad (4)$$

На основании известных свойств числовых рядов получаем, что ряды (3) сходятся по крайней мере при тех значениях x , при которых сходятся

¹⁾ При этом полагаем: $0! = 1! = 1$.

²⁾ Нам пока известен ряд геометрической прогрессии и пять основных разложений.

каждый из рядов (1) и (2). Другими словами, ряды (3) сходятся по крайней мере в меньшем из интервалов сходимости рядов (1), (2), и их суммы соответственно равны $f(x) \pm \varphi(x)$:

$$\sum_{n=0}^{\infty} a_n x^n \pm \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n \pm b_n) x^n.$$

Ряд (4) в силу абсолютной сходимости рядов (1), (2) в их интервалах сходимости сходится по крайней мере в меньшем из интервалов сходимости рядов (1) и (2), и его сумма равна произведению $f(x) \cdot \varphi(x)$:

$$\begin{aligned} \sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} b_n x^n &= (a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots) \times \\ &\times (b_0 + b_1 x + b_2 x^2 + \dots + b_n x^n + \dots) = a_0 b_0 + (a_0 b_1 + a_1 b_0) x + \\ &+ (a_0 b_2 + a_1 b_1 + a_2 b_0) x^2 + \dots + (a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0) x^n + \dots \end{aligned}$$

Не останавливаясь на этом, отметим только, что в результате рассматриваемых действий над рядами (1) и (2) может получиться ряд, который будет сходиться в интервале, содержащем внутри меньший из интервалов сходимости данных рядов (1) и (2). Но по крайней мере в меньшем из интервалов $(-R, R)$ и $(-R_1, R_1)$ сходимость ряда всегда имеет место.

Пример. Пусть $y = a_1 x + a_2 x^2 + a_3 x^3 + \dots + a_n x^n + \dots$ ($a_0 = 0$), причем ряд сходится в некоторой окрестности $x = 0$. Тогда в этой окрестности

$$\begin{aligned} y^2 &= (a_1 x + a_2 x^2 + a_3 x^3 + \dots)(a_1 x + a_2 x^2 + a_3 x^3 + \dots) = \\ &= a_1^2 x^2 + (a_1 a_2 + a_2 a_1) x^3 + (a_1 a_3 + a_2 a_2 + a_3 a_1) x^4 + \dots = \\ &= a_1^2 x^2 + 2a_1 a_2 x^3 + (a_2^2 + 2a_1 a_3) x^4 + \dots, \\ y^3 &= (a_1^2 x^2 + 2a_1 a_2 x^3 + \dots)(a_1 x + a_2 x^2 + a_3 x^3 + \dots) = \\ &= a_1^3 x^3 + (a_1^2 a_2 + 2a_1^2 a_2) x^4 + \dots = \\ &= a_1^3 x^3 + 3a_1^2 a_2 x^4 + \dots \\ &\dots \dots \dots \end{aligned}$$

Разложение в степенной ряд y^n (n — любое натуральное число) начнется с члена, содержащего x^n .

2) Подстановка ряда в ряд.

Теорема. Если функция $f(y)$ разлагается в степенной ряд, сходящийся в некоторой окрестности точки $y = 0$, $f(y) = c_0 + c_1 y + c_2 y^2 + \dots + c_n y^n + \dots$, и функция $y = \varphi(x)$ разлагается в степенной ряд по степеням x , сходящийся в некоторой окрестности $x = 0$, $\varphi(x) = a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$ ($\varphi(0) = 0$), то сложная функция $f[\varphi(x)]$ в некоторой окрестности $x = 0$ разлагается в степенной ряд, причем это разложение можно получить, группируя члены с одинаковыми степенями x в разложении¹⁾

$$f[\varphi(x)] = c_0 + c_1 \sum_{n=1}^{\infty} a_n x^n + c_2 \left(\sum_{n=1}^{\infty} a_n x^n \right)^2 + \dots + c_n \left(\sum_{n=1}^{\infty} a_n x^n \right)^n + \dots$$

Доказательства этой теоремы мы не приводим.

¹⁾ Заметим, что слагаемых, содержащих каждую фиксированную степень x , будет лишь конечное число.

Например, разложим $\ln \cos x$ в ряд по степеням x .
 Решение. $\ln \cos x = \ln [1 + (\cos x - 1)] = \ln (1 + y)$, где $y = \cos x - 1$. Но так как

$$\ln (1 + y) = y - \frac{1}{2} y^2 + \frac{1}{3} y^3 - \frac{1}{4} y^4 + \dots \quad (|y| < 1)$$

и

$$y = \cos x - 1 = -\frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \dots \quad (|x| < +\infty),$$

то, следовательно,

$$\begin{aligned} \ln \cos x &= \left(-\frac{1}{2} x^2 + \frac{1}{24} x^4 - \dots \right) - \frac{1}{2} \left(-\frac{1}{2} x^2 + \frac{1}{24} x^4 - \dots \right)^2 + \dots = \\ &= -\frac{1}{2} x^2 - \frac{1}{12} x^4 + \dots \end{aligned}$$

в некоторой окрестности $x=0$.

3) **Деление степенных рядов.** Пусть в некоторой окрестности $x=0$ сходятся ряды $a_0 + a_1 x + \dots + a_n x^n + \dots$ и $b_0 + b_1 x + b_2 x^2 + \dots + b_n x^n + \dots$, суммы которых соответственно равны $f(x)$ и $\varphi(x)$, причем $b_0 = \varphi(0) \neq 0$. Обозначим через $F(x)$ функцию, равную отношению сумм двух данных рядов:

$$F(x) = \frac{a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots}{b_0 + b_1 x + b_2 x^2 + \dots + b_n x^n + \dots} = \frac{f(x)}{\varphi(x)}.$$

Используя сформулированную выше теорему о подстановке ряда в ряд, докажем, что существует окрестность $x=0$, в которой функция $F(x)$ разлагается в степенной ряд, называемый *рядом частным двух данных рядов*.

Действительно,

$$F(x) = \frac{a_0 + a_1 x + \dots + a_n x^n + \dots}{b_0 + b_1 x + \dots + b_n x^n + \dots} = \frac{f(x)}{\varphi(x)} = \frac{f(x)}{b_0 + \psi(x)},$$

где $\psi(x) = b_1 x + b_2 x^2 + \dots + b_n x^n + \dots$ и $\psi(0) = 0$.

Или, что то же,

$$\begin{aligned} \frac{a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots}{b_0 + b_1 x + b_2 x^2 + \dots + b_n x^n + \dots} &= \frac{f(x)}{b_0} \cdot \frac{1}{1 + \frac{\psi(x)}{b_0}} = \\ &= \frac{f(x)}{b_0} \left[1 - \frac{\psi(x)}{b_0} + \left(\frac{\psi(x)}{b_0} \right)^2 - \dots \right]. \end{aligned}$$

Последнее равенство записано по формуле суммы геометрической прогрессии, знаменатель которой $q = -\frac{\psi(x)}{b_0}$. Такая запись возможна, так как $\psi(0) = 0$ и, следовательно, существует окрестность точки $x=0$, в которой имеет место неравенство $|q| = \left| \frac{\psi(x)}{b_0} \right| < 1$. Учитывая, что $\psi(x)$ — сумма степенного ряда, сходящегося в некоторой окрестности $x=0$ (на основании теоремы о подстановке ряда в ряд), заключаем, что выражение в квадратных скобках можно разложить в степенной ряд, сходящийся в некоторой окрестности нуля. Умножая этот ряд на ряд для функции $\frac{f(x)}{b_0}$, тоже сходящийся в какой-то окрестности $x=0$, получим степенной ряд, сходящийся в некоторой окрестности нуля.

Таким образом, доказана следующая теорема: частное двух степенных рядов, сходящихся в некоторой окрестности нуля, при условии, что свободный член ряда в знаменателе отличен от нуля, является степенным рядом, сходящимся в некоторой окрестности нуля.

Заметим, что в общем случае о величине этой окрестности или, что то же, о длине интервала сходимости частного двух степенных рядов ничего сказать нельзя.

Практически деление степенных рядов производят следующим образом: сначала записывают частное двух степенных рядов в виде степенного ряда, пока с неопределенными коэффициентами:

$$\frac{a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots}{b_0 + b_1x + b_2x^2 + \dots + b_nx^n + \dots} = c_0 + c_1x + c_2x^2 + \dots + c_nx^n + \dots$$

Затем, для определения коэффициентов c_n ($n=0, 1, \dots$) используют равенство

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots = (b_0 + b_1x + b_2x^2 + \dots + b_nx^n + \dots) \times \\ \times (c_0 + c_1x + c_2x^2 + \dots + c_nx^n + \dots),$$

из которого в результате умножения рядов и сравнения коэффициентов при одинаковых степенях x получают:

$$\begin{aligned} a_0 &= b_0c_0, \\ a_1 &= b_0c_1 + b_1c_0, \\ a_2 &= b_0c_2 + b_1c_1 + b_2c_0, \\ &\vdots \\ &\vdots \end{aligned}$$

Из этих равенств последовательно определяют c_0 (так как $b_0 \neq 0$), c_1 , c_2 и т. д.

Не доказывая, заметим, что деление степенных рядов можно производить так же, как и деление многочленов, расположенных по возрастающим степеням x .

Пример. Написать три члена разложения $\operatorname{tg} x$ в ряд по степеням x .

Решение: $\operatorname{tg} x = \frac{\sin x}{\cos x}$. Как известно,

$$\begin{aligned} \sin x &= x - \frac{x^3}{6} + \frac{x^5}{120} - \dots \quad (|x| < +\infty), \\ \cos x &= 1 - \frac{x^2}{2} + \frac{x^4}{24} - \dots \quad (|x| < +\infty). \end{aligned}$$

Тогда, производя деление рядов, как деление многочленов, получим:

$$\begin{array}{r|l} x - \frac{x^3}{6} + \frac{x^5}{120} - \dots & 1 - \frac{x^2}{2} + \frac{x^4}{24} - \dots \\ x - \frac{x^3}{2} + \frac{x^5}{24} - \dots & x + \frac{x^3}{3} + \frac{2}{15}x^5 + \dots \\ \hline & \frac{x^3}{3} - \frac{x^5}{30} + \dots \\ & - \frac{x^3}{3} + \frac{x^5}{6} + \dots \\ \hline & \frac{2}{15}x^5 + \dots \end{array}$$

т. е. $\operatorname{tg} x = x + \frac{x^3}{3} + \frac{12x^5}{15} + \dots$ в некоторой окрестности $x=0$.

§ 13. Приложения степенных рядов

Степенные ряды находят применение в различных вопросах как приближенных вычислений, так и в приближенных методах анализа. Укажем на некоторые приложения степенных рядов: с помощью степенных рядов составляют, например, таблицы значений функций, в частности, приближенно, с любой заданной точностью, вычисляют значения таких констант, как π , e и др.; с помощью степенных рядов вычисляют приближенные значения определенных интегралов (об этом будет сказано ниже) и т. д.

Как пример применения степенных рядов в приближенных методах анализа рассмотрим решение дифференциальных уравнений с помощью степенных рядов.

Решение дифференциальных уравнений с помощью степенных рядов. Рассмотрим линейное дифференциальное уравнение n -го порядка вида:

$$y^{(n)} + f_1(x) y^{(n-1)} + \dots + f_{n-1}(x) y' + f_n(x) y = F(x) \quad (1)$$

(с коэффициентом при старшей производной, равным единице).

Теорема. Если коэффициенты и правая часть дифференциального уравнения (1) разлагаются в степенные ряды по степеням $x - a$, сходящиеся в некоторой окрестности $x = a$, то решение этого уравнения, удовлетворяющее начальным условиям:

$$y(a) = y_0, \quad y'(a) = y_1, \quad \dots, \quad y^{(n-1)}(a) = y_{n-1}$$

($y_0, y_1, \dots, y_{n-1}$ — произвольно заданные числа), разлагается в степенной ряд по степеням $x - a$, сходящийся по крайней мере в меньшем из интервалов сходимости рядов для коэффициентов и правой части дифференциального уравнения.

Доказательство этой теоремы мы опускаем.

Практически решение дифференциального уравнения в виде степенного ряда можно получить двумя способами: сравнением коэффициентов и последовательным дифференцированием.

Способ сравнения коэффициентов заключается в следующем: сначала записывают решение в виде степенного ряда с неопределенными коэффициентами:

$$y = a_0 + a_1(x - a) + a_2(x - a)^2 + \dots + a_n(x - a)^n + \dots$$

Затем из начальных условий определяют значения коэффициентов $a_0, a_1, \dots, a_{n-1}$. После этого подставляют в дифференциальное уравнение вместо y и производных соответствующие степенные ряды, а также вместо коэффициентов и правой части записывают их разложения в степенные ряды по степеням $x - a$ и производят действия над рядами. Затем сравнивают коэффициенты при одинаковых степенях $x - a$. Из полученных таким путем уравнений определяют коэффициенты ряда.

Пример. Найти решение дифференциального уравнения $y'' - xy = 0$, удовлетворяющее начальным условиям $y(0) = 1, y'(0) = 0$.

Записываем искомое решение в виде ряда: $y = a_0 + a_1x + a_2x^2 + \dots$ и из начальных условий определяем a_0 и a_1 :

$$\begin{aligned} y(0) &= 1 = a_0, & a_0 &= 1; \\ y'(0) &= 0 = a_1, & a_1 &= 0. \end{aligned}$$

Затем $y = 1 + a_2x^2 + a_3x^3 + \dots$ подставляем в уравнение

$$2a_2 + 2 \cdot 3a_3x + 3 \cdot 4a_4x^2 + \dots - (x + a_2x^3 + a_3x^4 + \dots) = 0,$$

откуда, сравнивая коэффициенты, получаем:

$$\begin{aligned} 2a_2 &= 0, & a_2 &= 0; \\ 2 \cdot 3a_3 &= 1, & a_3 &= \frac{1}{2 \cdot 3}; \\ 3 \cdot 4a_4 &= 0, & a_4 &= 0; \\ 4 \cdot 5a_5 &= a_2, & a_5 &= 0; \\ 5 \cdot 6a_6 &= a_3, & a_6 &= \frac{1}{2 \cdot 3 \cdot 5 \cdot 6} \cdot \\ & \dots & & \end{aligned}$$

(В этом примере нетрудно подметить закон образования коэффициентов, предоставляем это сделать читателю.)

Полученный ряд, как это следует из теоремы, сходится при всех значениях x .

При нахождении решения дифференциального уравнения *способом последовательного дифференцирования* разлагают искомое решение в ряд Тейлора¹⁾:

$$y = y(a) + \frac{y'(a)}{1!}(x-a) + \frac{y''(a)}{2!}(x-a)^2 + \dots + \frac{y^{(n)}(a)}{n!}(x-a)^n + \dots$$

При этом первые n коэффициенты $y(a)$, $y'(a)$, ..., $y^{(n-1)}(a)$ заданы начальными условиями. Подставляя в уравнение (1) $x=a$, определяют $y^{(n)}(a)$. Последовательно дифференцируя уравнение (1) и подставляя $x=a$, определяют коэффициенты ряда искомого решения.

Пример. Найти решение дифференциального уравнения

$$y'' - xy = 0. \quad (2)$$

При начальных условиях $y(0) = 1$, $y'(0) = 0$. (Это уравнение уже решалось методом сравнения коэффициентов.)

Запишем решение уравнения в виде:

$$y = y(0) + \frac{y'(0)}{1!}x + \frac{y''(0)}{2!}x^2 + \frac{y'''(0)}{3!}x^3 + \dots$$

По условию $y(0) = 1$, $y'(0) = 0$.

Подставляя в дифференциальное уравнение $x=0$, получим:

$$y''(0) = 0.$$

Будем теперь последовательно дифференцировать уравнение (2) и подставлять значение $x=0$, тогда

$$\begin{array}{ll} y''' = y + xy', & y'''(0) = 1; \\ y^{(IV)} = 2y' + xy'', & y^{(IV)}(0) = 0; \\ y^{(V)} = 3y'' + xy''', & y^{(V)}(0) = 0; \\ y^{(VI)} = 4y''' + xy^{(IV)}, & y^{(VI)}(0) = 4; \\ \dots & \dots \end{array}$$

И, наконец, подставляя в выражение y полученные значения производных, запишем:

$$y = 1 + \frac{x^3}{2 \cdot 3} + \frac{x^6}{2 \cdot 3 \cdot 5 \cdot 6} + \dots$$

Если коэффициенты и правая часть дифференциального уравнения (1) разлагаются в степенные ряды по степеням $x-a$, сходящиеся в некоторой окрестности $x=a$, то, как уже говорилось выше, любое в окрестности $x=a$ решение уравнения (1) разлагается в степенной ряд по степеням $x-a$, сходящийся по крайней мере в меньшем из интервалов сходимости рядов для коэффициентов и правой части дифференциального уравнения (1).

Таким образом любое решение в окрестности $x=a$ можно записать в виде степенного ряда по степеням $x-a$. При этом все коэффициенты ряда-решения будут выражаться через первые коэффициенты этого ряда, которые могут принимать любые действительные значения, т. е. являются произвольными постоянными общего решения дифференциального уравнения.

Биномиальный ряд. Рассмотрим дифференциальное уравнение

$$(1+x)y' - ay = 0 \quad (2')$$

¹⁾ Как следует из теоремы, искомое решение разлагается в ряд Тейлора по степеням $x-a$.

(α — любое действительное число) и найдем его решение при начальном условии $y(0) = 1$.

Прежде всего отметим, что решением этого уравнения с разделяющимися переменными, удовлетворяющим заданному начальному условию, будет функция $y = (1+x)^\alpha$. (Получить это предоставляем читателю.)

Так как заданное дифференциальное уравнение (2') можно записать в виде: $y' - \frac{\alpha}{1+x} y = 0$ и коэффициент $\frac{\alpha}{1+x}$ разлагается в окрестности $x=0$ в степенной ряд (радиус сходимости этого ряда, как известно, равен 1), то решение уравнения (2') можно искать в виде степенного ряда, при этом радиус сходимости ряда-решения будет равен единице. Тогда ряд-решение в силу единственности решения уравнения, удовлетворяющего начальному условию, и будет разложением в степенной ряд функции $(1+x)^\alpha$.

Итак, запишем искомое решение в виде ряда:

$$y = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$$

Подчинив y начальному условию, найдем $y(0) = a_0 = 1$. Подставляя y в дифференциальное уравнение, получим:

$$(1+x)[a_1 + 2a_2 x + 3a_3 x^2 + \dots + (n-1)a_{n-1} x^{n-2} + na_n x^{n-1} + \dots] = \\ = \alpha(1 + a_1 x + a_2 x^2 + a_3 x^3 + \dots + a_n x^n + \dots)$$

или, расположив левую часть равенства по степеням x , имеем

$$a_1 + (2a_2 + a_1)x + (3a_3 + 2a_2)x^2 + \dots + [(n+1)a_{n+1} + na_n]x^n + \dots = \\ = \alpha(1 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots).$$

Сравнивая коэффициенты при одинаковых степенях x , получаем:

$$\begin{aligned} a_1 &= \alpha, \\ 2a_2 + a_1 &= \alpha a_1, \\ 3a_3 + 2a_2 &= \alpha a_2, \\ &\dots \dots \dots \\ (n+1)a_{n+1} + na_n &= \alpha a_n, \\ &\dots \dots \dots \end{aligned}$$

Откуда,

$$\begin{aligned} a_1 &= \alpha, \\ a_2 &= \frac{\alpha-1}{2} a_1 = \frac{\alpha(\alpha-1)}{1 \cdot 2}, \\ a_3 &= \frac{\alpha-2}{3} a_2 = \frac{\alpha(\alpha-1)(\alpha-2)}{1 \cdot 2 \cdot 3}, \\ &\dots \dots \dots \\ a_{n+1} &= \frac{\alpha-n}{n+1} a_n, \\ &\dots \dots \dots \end{aligned}$$

Таким образом,

$$y = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1 \cdot 2} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{1 \cdot 2 \cdot 3} x^3 + \dots + \\ + \frac{\alpha(\alpha-1) \dots [\alpha-(n-1)]}{1 \cdot 2 \cdot 3 \dots n} x^n + \dots \quad (|x| < 1)$$

или, что то же,

$$y = \sum_{n=0}^{\infty} \frac{\alpha(\alpha-1)(\alpha-2) \dots [\alpha-(n-1)]}{n!} x^n \quad (|x| < 1). \quad (3)$$

Заметим, что сходимость ряда (3) при $|x| < 1$ можно получить также, используя признак Даламбера. Действительно,

$$\lim_{n \rightarrow \infty} \frac{|\alpha(\alpha-1)(\alpha-2)\dots[\alpha-(n-1)](\alpha-n)!| |x|^{n+1} \cdot n!}{(n+1)! |\alpha(\alpha-1)(\alpha-2)\dots[\alpha-(n-1)]| |x|^n} = |x| \lim_{n \rightarrow \infty} \frac{|\alpha-n|}{(n+1)} = |x|,$$

при $|x| < 1$ ряд сходится,
при $|x| > 1$ ряд расходится, т. е. $R=1$.

Следовательно, при $|x| < 1$ справедливо равенство

$$(1+x)^\alpha = 1 + \frac{\alpha}{1}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \dots + \frac{\alpha(\alpha-1)(\alpha-2)\dots[\alpha-(n-1)]}{n!}x^n + \dots, \quad (|x| < 1). \quad (4)$$

Ряд (4) функции $(1+x)^\alpha$ называется *биномиальным рядом*. При α целом положительном равенство (4) превращается в известную формулу бинома Ньютона, так как в этом случае все коэффициенты, начиная с некоторого, равны нулю.

Пример. Разложить функцию $f(x) = \arcsin x$ в ряд по степеням x .

Для решений этой задачи воспользуемся равенством $\arcsin x = \int_0^x \frac{dx}{\sqrt{1-x^2}}$ и биномиальным рядом. Действительно,

$$\begin{aligned} \frac{1}{\sqrt{1+x}} &= (1+x)^{-\frac{1}{2}} = 1 - \frac{1}{2}x + \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{1 \cdot 2}x^2 + \\ &+ \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)}{1 \cdot 2 \cdot 3}x^3 + \dots = 1 - \frac{1}{2}x + \frac{1 \cdot 3}{2 \cdot 4}x^2 - \\ &- \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^3 + \dots + (-1)^n \frac{1 \cdot 3 \cdot 5 \dots (2n-1)}{2 \cdot 4 \cdot 6 \dots 2n}x^n + \dots \end{aligned}$$

при $|x| < 1$.

Заменяя в полученном равенстве x на $-x^2$, имеем:

$$\frac{1}{\sqrt{1-x^2}} = 1 + \frac{1}{2}x^2 + \frac{1 \cdot 3}{2 \cdot 4}x^4 + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^6 + \dots + \frac{1 \cdot 3 \cdot 5 \dots (2n-1)}{2 \cdot 4 \cdot 6 \dots 2n}x^{2n} + \dots$$

при $|-x^2| < 1$ или, что то же, $|x| < 1$.

Почленно интегрируя этот ряд в пределах от 0 до x , $|x| < 1$, получим:

$$\arcsin x = x + \frac{1}{2} \cdot \frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \frac{x^5}{5} + \dots + \frac{1 \cdot 3 \cdot 5 \dots (2n-1)}{2 \cdot 4 \cdot 6 \dots 2n} \cdot \frac{x^{2n+1}}{2n+1} + \dots$$

при $|x| < 1$.

Положив $x = \frac{1}{2}$, определим:

$$\frac{\pi}{6} = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3 \cdot 2^3} + \frac{1 \cdot 3}{2 \cdot 4} \cdot \frac{1}{5 \cdot 2^5} + \dots$$

Это равенство дает возможность приближенно вычислить число π .

Вычисление интегралов с помощью степенных рядов. В § 14 гл. VI были рассмотрены способы прямоугольников, трапеций, парабол для приближенного интегрирования. Здесь мы рассмотрим на примере интеграла

$\frac{1}{2} \int_0^1 e^{-x^2} dx$ способ вычисления интегралов с помощью степенных рядов.

Вычислим интеграл $\int_0^{\frac{1}{2}} e^{-x^2} dx$. Для этого разложим подинтегральную функцию в степенной ряд. Так как $e^t = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \dots + \frac{t^n}{n!} + \dots$ при любом t , то, подставляя $-x^2$ вместо t , получим:

$$e^{-x^2} = 1 - \frac{x^2}{1!} + \frac{x^4}{2!} - \frac{x^6}{3!} + \dots + (-1)^n \frac{x^{2n}}{n!} + \dots \quad (|x| < +\infty).$$

Почленно интегрируя, найдем:

$$\begin{aligned} \int_0^{\frac{1}{2}} e^{-x^2} dx &= \left[x - \frac{x^3}{3} + \frac{1}{2!} \frac{x^5}{5} - \frac{1}{3!} \frac{x^7}{7} + \dots \right]_0^{\frac{1}{2}} = \\ &= \frac{1}{2} - \frac{1}{3} \cdot \frac{1}{2^3} + \frac{1}{2! \cdot 5} \cdot \frac{1}{2^5} - \frac{1}{3! \cdot 7} \cdot \frac{1}{2^7} + \dots \end{aligned}$$

Получился знакочередующийся ряд, удовлетворяющий условиям признака Лейбница, следовательно, ошибка при замене суммы ряда его частной суммой по абсолютной величине меньше абсолютной величины первого из отброшенных его членов. В частности, положив, что интеграл равен сумме первых двух слагаемых, мы делаем ошибку, меньшую $\frac{1}{2! \cdot 5} \cdot \frac{1}{2^5} < 0,003$. Отсюда следует, что, ограничиваясь только двумя слагаемыми, мы получаем приближенное значение интеграла с точностью до 0,01:

$$\int_0^{\frac{1}{2}} e^{-x^2} dx \approx \frac{1}{2} - \frac{1}{3} \cdot \frac{1}{2^3} \quad \text{или} \quad \int_0^{\frac{1}{2}} e^{-x^2} dx \approx 0,46.$$

Раздел 4

НЕКОТОРЫЕ ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ КОМПЛЕКСНОЙ ПЕРЕМЕННОЙ

§ 14. Показательная, тригонометрические и гиперболические функции

В § 11 этой главы были получены разложения функций e^x , $\sin x$, $\cos x$, $\operatorname{sh} x$, $\operatorname{ch} x$ в степенные ряды, сходящиеся при всех значениях x .

$$\begin{aligned} e^x &= \sum_{n=1}^{\infty} \frac{x^n}{n!}, \quad \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \\ \cos x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad \operatorname{sh} x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}, \\ \operatorname{ch} x &= \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}. \end{aligned}$$

Если в правых частях записанных равенств под x будем понимать комплексные числа, обозначив их через z , то получим ряды по степе-

ням z , сходящиеся, и притом абсолютно (как это следует из теоремы Абеля), при всех комплексных значениях z , т. е. во всей комплексной плоскости. Суммы рядов

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad (1')$$

$$\sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}, \quad (2')$$

$$\sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}, \quad (3')$$

$$\sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!}, \quad (4')$$

$$\sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} \quad (5')$$

являются функциями комплексной переменной z , определенными во всей плоскости z .

Естественно сумму ряда (1') назвать *показательной функцией комплексной переменной z* и обозначить через e^z , так как для действительных значений z , при $z=x$, сумма ряда (1') равна e^x . Аналогично суммы других рядов соответственно называют *синусом, косинусом, синусом гиперболическим и косинусом гиперболическим комплексной переменной z* и соответственно обозначают: $\sin z$, $\cos z$, $\operatorname{sh} z$, $\operatorname{ch} z$.

Итак, по определению для любого комплексного числа z

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad (1)$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}, \quad (2)$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}, \quad (3)$$

$$\operatorname{sh} z = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!}, \quad (4)$$

$$\operatorname{ch} z = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \dots = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!}. \quad (5)$$

Отметим еще раз, что для действительных значений z эти функции получают уже известные нам значения.

Изучение свойств рассматриваемых функций комплексной переменной z должно производиться на основании принятых определений, т. е. исходя из свойств рядов. Эти ряды дают возможность вычислить значение каждой из этих функций при любом определенном значении z , но обычно при вычислении значений этих функций используют их свойства, которые будут получены дальше.

Прежде всего, непосредственно из определения этих функций следует, что $\sin z$ и $\operatorname{sh} z$ — нечетные функции ($f(-z) = -f(z)$), а $\cos z$ и $\operatorname{ch} z$ — четные функции ($f(-z) = f(z)$).

Формула Эйлера. Подставляя в равенство (1) iz вместо z и производя надлежащую группировку членов ряда, что возможно в абсолютно сходящемся ряде, получим:

$$\begin{aligned} e^{iz} &= 1 + iz + \frac{i^2 z^2}{2!} + \frac{i^3 z^3}{3!} + \frac{i^4 z^4}{4!} + \frac{i^5 z^5}{5!} + \dots = \\ &= 1 + iz - \frac{z^2}{2!} - i \frac{z^3}{3!} + \frac{z^4}{4!} + i \frac{z^5}{5!} - \dots = \\ &= \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots\right) + i \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots\right) = \cos z + i \sin z. \end{aligned}$$

Равенство

$$e^{iz} = \cos z + i \sin z \quad (6)$$

называется *формулой Эйлера*. С помощью формулы Эйлера легко получить выражение $\sin z$ и $\cos z$ через показательную функцию. Действительно, подставляя в формулу Эйлера $-z$ вместо z и учитывая четность $\cos z$ и нечетность $\sin z$, получим:

$$e^{-iz} = \cos z - i \sin z. \quad (7)$$

Почленно складывая и вычитая равенства (6) и (7), найдем:

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}. \quad (8)$$

Эти формулы также называют *формулами Эйлера*.

Заметим, что формула Эйлера верна при любом комплексном значении z , в частности, при любом действительном значении z , $z = x$, она принимает вид:

$$e^{ix} = \cos x + i \sin x.$$

Полученное равенство в теории дифференциальных уравнений фигурировало как определение чисто мнимой степени числа e .

Как следует из формул (8), функции $\sin z$ и $\cos z$ выражаются линейно через показательную функцию. Такая связь стала возможной только при выходе в комплексную плоскость.

Из формул Эйлера легко получить, что свойство ограниченности модулей функций $\sin z$ и $\cos z$, которое имеет место в действительной области, не сохраняется в комплексной области.

Действительно, при $z = i$ $\cos i = \frac{e^{-1} + e^1}{2} > 1$, при $z = iy$ $\cos iy = \frac{e^{-y} + e^y}{2}$ неограниченно растёт, когда $y \rightarrow \infty$.

Рассмотрим еще одно приложение формулы Эйлера. Как известно, правая часть равенства $z = r(\cos \varphi + i \sin \varphi)$, где r — модуль комплексного числа z , а φ — аргумент, называется *тригонометрической формой комплексного числа* z . Учитывая формулу Эйлера, можно записать: $z = re^{i\varphi}$. Выражение $re^{i\varphi}$ называется *показательной формой комплексного числа* z .

Связь между показательной и гиперболическими функциями. Известная связь между этими функциями в области действительных значений аргумента сохраняется и для комплексных значений аргумента.

В самом деле, положив в равенстве

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots + \frac{z^n}{n!} + \dots$$

— z вместо z , получим:

$$e^{-z} = 1 - \frac{z}{1!} + \frac{z^2}{2!} - \frac{z^3}{3!} + \frac{z^4}{4!} + \dots + (-1)^n \frac{z^n}{n!} + \dots,$$

откуда

$$\frac{e^z - e^{-z}}{2} = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} = \operatorname{sh} z,$$

$$\frac{e^z + e^{-z}}{2} = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} = \operatorname{ch} z.$$

И, следовательно, мы выразили гиперболические функции через показательную функцию:

$$\operatorname{sh} z = \frac{e^z - e^{-z}}{2}, \quad \operatorname{ch} z = \frac{e^z + e^{-z}}{2}. \quad (9)$$

Связь между тригонометрическими и гиперболическими функциями. Связь между этими функциями устанавливается с помощью их выражений через показательную функцию, т. е. с помощью формул (8) и (9).

Действительно,

$$\operatorname{ch} iz = \frac{e^{iz} + e^{-iz}}{2} = \cos z,$$

$$\operatorname{sh} iz = \frac{e^{iz} - e^{-iz}}{2} = i \frac{e^{iz} - e^{-iz}}{2i} = i \sin z,$$

откуда следует

$$\left. \begin{aligned} \operatorname{ch} iz &= \cos z, \\ \operatorname{sh} iz &= i \sin z. \end{aligned} \right\} \quad (10)$$

Подставляя в эти равенства iz вместо z , получим:

$$\begin{aligned} \cos iz &= \operatorname{ch} (-z) = \operatorname{ch} z, \\ i \sin iz &= \operatorname{sh} (-z) = -\operatorname{sh} z \end{aligned}$$

или

$$\sin iz = i \operatorname{sh} z$$

и, значит,

$$\left. \begin{aligned} \cos iz &= \operatorname{ch} z, \\ \sin iz &= i \operatorname{sh} z. \end{aligned} \right\} \quad (11)$$

Теорема сложения для показательной функции. Теорема сложения является основой для изучения свойств показательной, а также тригонометрических и гиперболических функций. Ее можно сформулировать так: для любых двух комплексных чисел z_1 и z_2 выполняется равенство

$$e^{z_1 + z_2} = e^{z_1} \cdot e^{z_2}.$$

Как известно, для действительных значений z_1 и z_2 это равенство справедливо.

Доказательство. Так как ряды $e^{z_1} = \sum_{n=0}^{\infty} \frac{z_1^n}{n!}$, $e^{z_2} = \sum_{n=0}^{\infty} \frac{z_2^n}{n!}$ абсолютно сходятся, то их можно почленно перемножать.

Умножая ряды

$$\begin{aligned} e^{z_1} &= 1 + \frac{z_1}{1!} + \frac{z_1^2}{2!} + \frac{z_1^3}{3!} + \dots + \frac{z_1^n}{n!} + \dots, \\ e^{z_2} &= 1 + \frac{z_2}{1!} + \frac{z_2^2}{2!} + \frac{z_2^3}{3!} + \dots + \frac{z_2^n}{n!} + \dots, \end{aligned}$$

получим:

$$\begin{aligned}
 e^{z_1} \cdot e^{z_2} &= 1 + \left(\frac{z_1 + z_2}{1!} \right) + \left(\frac{z_1^2}{2!} + \frac{z_1 z_2}{1!1!} + \frac{z_2^2}{2!} \right) + \\
 &+ \left(\frac{z_1^3}{3!} + \frac{z_1^2 z_2}{2!1!} + \frac{z_1 z_2^2}{1!2!} + \frac{z_2^3}{3!} \right) + \dots + \left(\frac{z_1^n}{n!} + \frac{z_1^{n-1} z_2}{(n-1)!1!} + \dots + \frac{z_2^n}{n!} \right) + \dots = \\
 &= 1 + (z_1 + z_2) + \frac{1}{2!} (z_1^2 + 2z_1 z_2 + z_2^2) + \\
 &+ \frac{1}{3!} (z_1^3 + 3z_1^2 z_2 + 3z_1 z_2^2 + z_2^3) + \dots + \\
 &+ \frac{1}{n!} (z_1^n + n z_1^{n-1} z_2 + \dots + z_2^n) + \dots = 1 + (z_1 + z_2) + \\
 &+ \frac{(z_1 + z_2)^2}{2!} + \frac{(z_1 + z_2)^3}{3!} + \dots + \frac{(z_1 + z_2)^n}{n!} + \dots = e^{z_1 + z_2}, \\
 e^{z_1} \cdot e^{z_2} &= e^{z_1 + z_2}.
 \end{aligned} \tag{12}$$

Из теоремы сложения и формулы Эйлера (6) можно получить ряд свойств рассматриваемых функций.

Для этого сначала найдем действительную и мнимую части показательной функции e^z .

$$e^z = e^{x+iy} = e^x \cdot e^{iy} = e^x (\cos y + i \sin y).$$

Следовательно,

$$\begin{aligned}
 \operatorname{Re}(e^z) &= e^x \cos y, \\
 \operatorname{Im}(e^z) &= e^x \sin y
 \end{aligned}$$

и модуль e^z запишется в виде $|e^z| = \sqrt{e^{2x}(\cos^2 y + \sin^2 y)} = e^x$ или, что то же,

$$|e^z| = e^{\operatorname{Re}(z)}. \tag{13}$$

Например, найдем $|e^{iz}|$ и $|e^{z^2}|$. Так как $e^{iz} = e^{i(x+iy)} = e^{-y+ix}$, то $|e^{iz}| = e^{-y}$.

Аналогично имеем $e^{z^2} = e^{(x+iy)^2} = e^{x^2-y^2+i2xy}$ и, следовательно, $|e^{z^2}| = e^{x^2-y^2}$.

Заметим, как следует из равенства $\operatorname{Im}(e^z) = e^x \sin y$, что функция e^z всегда принимает мнимые значения при условии $y \neq k\pi$ (k — любое целое число).

Если же $y = k\pi$ (k — любое целое число), то функция e^z принимает действительные значения, в частности, при $k=0$ получим действительные значения для z и для функции e^z .

Учитывая выражение для модуля e^z , $|e^z| = e^x$, а также тот факт, что показательная функция e^x при действительных значениях x в нуль не обращается, заключаем, что функция e^z нигде в нуль не обращается.

Показательная функция e^z является периодической функцией, ее период равен $2\pi i$. Действительно,

$$e^{z+2\pi i} = e^z \cdot e^{2\pi i} = e^z (\cos 2\pi + i \sin 2\pi) = e^z.$$

Из этого свойства показательной функции и формул (9) следует, что гиперболические функции $\operatorname{sh} z$ и $\operatorname{ch} z$ также являются периодическими функциями с периодом $2\pi i$. Последнее означает, что значения функций e^z , $\operatorname{sh} z$, $\operatorname{ch} z$ в точке z и в каждой точке, получающейся из z сдвигом по вертикали на 2π (или на $2\pi k$, k — целое число), между собой соответственно равны.

Функция e^{iz} имеет период, равный 2π , так как $e^{i(z+2\pi)} = e^{iz} \cdot e^{2\pi i} = e^{iz}$. И, следовательно, учитывая формулы (8), получим, что функции $\sin z$ и $\cos z$ имеют период, равный 2π , т. е. свойство, ранее известное для тригонометрических функций в действительной области, остается справедливым и в комплексной области.

И, наконец, что, перемножая равенства (6) и (7), получим:

$$(\cos z + i \sin z) \cdot (\cos z - i \sin z) = e^{iz} \cdot e^{-iz} = 1$$

или, что то же,

$$\cos^2 z + \sin^2 z = 1.$$

Таким образом, основное равенство для тригонометрических функций сохраняется и при комплексных значениях аргумента.

Используя связь между тригонометрическими и гиперболическими функциями и учитывая, что $\cos^2 iz + \sin^2 iz = 1$, получим: $\operatorname{ch}^2 z + (i \operatorname{sh} z)^2 = 1$ или, что то же, $\operatorname{ch}^2 z - \operatorname{sh}^2 z = 1$. Это равенство является основным соотношением для гиперболических функций.

Рассмотрим ошибку, которую допускают при отыскании модуля функции e^{iz} , когда рассуждают так:

$$e^{iz} = \cos z + i \sin z,$$

поэтому

$$|e^{iz}| = \sqrt{\cos^2 z + \sin^2 z} = 1$$

и получают результат, противоречащий уже известному:

$$|e^{iz}| = e^{-y}.$$

Действительно, сравнивая эти два равенства, получим, что при любом z , а следовательно, и при любом $y = I(z)$, справедливо соотношение $e^{-y} = 1$, что нелепо. Такая ошибка произошла за счет того, что «посчитали» функции $\cos z$ и $\sin z$ действительными при любых значениях z . На самом же деле функции $\cos z$ и $\sin z$ принимают действительные значения только при действительных значениях аргумента.

Это утверждение рекомендуется проверить самостоятельно, используя, например, формулы (8) или результаты следующего пункта.

Теоремы сложения для тригонометрических и гиперболических функций. Из теоремы сложения для показательной функции и формул Эйлера получаем:

$$\begin{aligned} \sin(z_1 + z_2) &= \frac{e^{i(z_1 + z_2)} - e^{-i(z_1 + z_2)}}{2i} = \frac{e^{iz_1} \cdot e^{iz_2} - e^{-iz_1} \cdot e^{-iz_2}}{2i} = \\ &= \frac{(\cos z_1 + i \sin z_1)(\cos z_2 + i \sin z_2) - (\cos z_1 - i \sin z_1)(\cos z_2 - i \sin z_2)}{2i} = \\ &= \sin z_1 \cos z_2 + \cos z_1 \sin z_2. \end{aligned}$$

Таким образом,

$$\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2. \quad (14)$$

Аналогично имеем:

$$\begin{aligned} \cos(z_1 + z_2) &= \frac{e^{i(z_1 + z_2)} + e^{-i(z_1 + z_2)}}{2} = \frac{e^{iz_1} \cdot e^{iz_2} + e^{-iz_1} \cdot e^{-iz_2}}{2} = \\ &= \frac{(\cos z_1 + i \sin z_1)(\cos z_2 + i \sin z_2) + (\cos z_1 - i \sin z_1)(\cos z_2 - i \sin z_2)}{2} = \\ &= \cos z_1 \cos z_2 - \sin z_1 \sin z_2, \end{aligned}$$

т. е.

$$\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2. \quad (15)$$

Равенства (14), (15) называются *теоремами сложения* соответственно для *синуса* и *косинуса*.

Заметим, что, используя равенства (15), (14), а также основное соотношение $\sin^2 z + \cos^2 z = 1$, можно показать, что основные формулы для тригонометрических функций, доказанные в действительной области, остаются верными и в комплексной области.

Подставляя в равенства (14) и (15) вместо z_1 и z_2 соответственно iz_1 и iz_2 и учитывая формулы (11), получим теоремы сложения для гиперболических функций:

$$\begin{aligned} \sin(iz_1 + iz_2) &= \sin i(z_1 + z_2) = \sin iz_1 \cos iz_2 + \cos iz_1 \sin iz_2, \\ i \operatorname{sh}(z_1 + z_2) &= i \operatorname{sh} z_1 \operatorname{ch} z_2 + i \operatorname{ch} z_1 \operatorname{sh} z_2 \end{aligned}$$

или

$$\operatorname{sh}(z_1 + z_2) = \operatorname{sh} z_1 \operatorname{ch} z_2 + \operatorname{ch} z_1 \operatorname{sh} z_2. \quad (16)$$

Аналогично получаем:

$$\cos(iz_1 + iz_2) = \cos i(z_1 + z_2) = \cos iz_1 \cos iz_2 - \sin iz_1 \sin iz_2$$

или

$$\operatorname{ch}(z_1 + z_2) = \operatorname{ch} z_1 \operatorname{ch} z_2 + \operatorname{sh} z_1 \operatorname{sh} z_2. \quad (17)$$

Равенства (16), (17) называются *теоремами сложения* соответственно для гиперболических синуса и косинуса.

Заметим, что из каждой формулы, справедливой для тригонометрических функций, можно получить соответствующую формулу для гиперболических функций, если использовать связь между тригонометрическими и гиперболическими функциями, т. е. равенства (10), (11).

В качестве примера на теоремы сложения найдем $\operatorname{Re}(\sin z)$, $\operatorname{Im}(\sin z)$ и $|\sin z|$.

Решение. Так как

$$\sin z = \sin(x + iy) = \sin x \cos iy + \cos x \sin iy = \sin x \operatorname{ch} y + i \cos x \operatorname{sh} y,$$

то

$$\operatorname{Re}(\sin z) = \sin x \operatorname{ch} y,$$

$$\operatorname{Im}(\sin z) = \cos x \operatorname{sh} y$$

и, следовательно,

$$\begin{aligned} |\sin z|^2 &= (\sin x \operatorname{ch} y)^2 + (\cos x \operatorname{sh} y)^2 = \sin^2 x \operatorname{ch}^2 y + \cos^2 x \operatorname{sh}^2 y \\ &= \sin^2 x (1 + \operatorname{sh}^2 y) + \cos^2 x \operatorname{sh}^2 y = \sin^2 x + \operatorname{sh}^2 y, \end{aligned}$$

т. е. $|\sin z| = \sqrt{\sin^2 x + \operatorname{sh}^2 y}$.

Из полученного равенства легко определить нули функции $\sin z$. Действительно, $\sin z = 0$ при условии, что $\sin x = 0$ и $\operatorname{sh} y = 0$, т. е. при $x = k\pi$ (k — целое число) и $y = 0$.

Окончательно $\sin z = 0$ только при $z = k\pi$ (k — целое).

Аналогично можно получить, что $\cos z = 0$ только при $z = \frac{\pi}{2} + k\pi$ (k — целое).

Учитывая выражение гиперболических функций через тригонометрические функции, можно также найти нули $\operatorname{sh} z$ и $\operatorname{ch} z$. Рекомендуется проделать это самостоятельно.

Определение тригонометрического и гиперболического тангенса комплексного аргумента. Названные функции обозначаются соответственно $\operatorname{tg} z$ и $\operatorname{th} z$ и определяются равенствами:

$$\begin{aligned} \operatorname{tg} z &= \frac{\sin z}{\cos z} = \frac{1}{i} \frac{e^{iz} - e^{-iz}}{e^{iz} + e^{-iz}}, \\ \operatorname{th} z &= \frac{\operatorname{sh} z}{\operatorname{ch} z} = \frac{e^z - e^{-z}}{e^z + e^{-z}} \end{aligned}$$

(при этом предполагается, что $\cos z \neq 0$ и $\operatorname{ch} z \neq 0$).

Нетрудно установить, что между этими функциями существует следующая связь:

$$\begin{aligned} \operatorname{tg}(iz) &= i \operatorname{th} z, \\ \operatorname{th}(iz) &= i \operatorname{tg} z. \end{aligned}$$

§ 15. Некоторые многозначные функции комплексного аргумента

Логарифмическая функция. Логарифмом числа z по основанию e (обозначается $\operatorname{Ln} z$) называется всякое число w , удовлетворяющее равенству $e^w = z$.

Из определения $\operatorname{Ln} z$ следует, что $z \neq 0$, так как показательная функция в нуль не обращается.

Запишем z в показательной форме, а w — в алгебраической форме: $z = re^{i\varphi}$; $w = u + iv$. Тогда по определению логарифма $e^{u+iv} = re^{i\varphi}$ или $e^u \cdot e^{iv} = re^{i\varphi}$.

Из равенства комплексных чисел следует равенство их модулей: $e^u = r$, откуда $u = \ln r$, и соотношение между аргументами: $v = \varphi + 2k\pi$ (k — любое целое число) и, значит,

$$w = \ln r + i(\varphi + 2k\pi).$$

Обратно, при любом целом k записанное выражение, как непосредственно видно, является значением логарифма числа z . Итак, $\text{Ln } z$ (при $z \neq 0$) имеет бесчисленное множество значений, причем все они получаются по формуле

$$\text{Ln } z = \ln |z| + i(\varphi + 2k\pi) \quad (k=0, \pm 1, \pm 2, \dots)$$

или, что то же, $\text{Ln } z = \ln |z| + i \text{Arg } z$.

Так как в области комплексных значений каждое число, не равное нулю, имеет бесчисленное множество значений логарифма, то $\text{Ln } z$ — бесконечнозначная функция, определенная во всей плоскости, кроме $z=0$.

Если рассматривать только действительные значения переменной z и функции, т. е. рассматривать логарифмическую функцию в действительной области, то она будет определена лишь для положительных значений аргумента (в этом случае $\varphi=0$) и однозначна для этих значений. Для отрицательных же значений аргумента логарифмическая функция в действительной области не существует (в этом случае $\varphi=\pi$). И мы получили известное свойство логарифмов в действительной области, которое коротко формулируется так: отрицательные числа логарифмов не имеют.

Обозначим через $\ln z = \ln |z| + i\varphi$, где $-\pi < \varphi \leq \pi$, и назовем это выражение *главным значением логарифма*, тогда

$$\text{Ln } z = \ln z + 2k\pi i \quad (k=0, \pm 1, \pm 2, \dots).$$

Пример 1.

$$\begin{aligned} \text{Ln } 1 &= \ln 1 + i(0 + 2k\pi) = 2k\pi i \\ \text{Ln } (-1) &= \ln 1 + i(\pi + 2k\pi) = (2k+1)\pi i \\ \text{Ln } (1+i) &= \ln \sqrt{2} + i\left(\frac{\pi}{4} + 2k\pi\right). \end{aligned}$$

Пример 2. Доказать основное свойство логарифмов

$$\text{Ln } (z_1 \cdot z_2) = \text{Ln } z_1 + \text{Ln } z_2.$$

Действительно, из тождества $e^{\text{Ln } z} = z$ получаем $e^{\text{Ln } z_1} = z_1$, $e^{\text{Ln } z_2} = z_2$. По теореме сложения показательной функции:

$$e^{\text{Ln } z_1 + \text{Ln } z_2} = e^{\text{Ln } z_1} \cdot e^{\text{Ln } z_2} = z_1 \cdot z_2.$$

Но $e^{\text{Ln } (z_1 \cdot z_2)} = z_1 \cdot z_2$. Из сравнения этих равенств получаем:

$$e^{\text{Ln } (z_1 \cdot z_2)} = e^{\text{Ln } z_1 + \text{Ln } z_2},$$

откуда $\text{Ln } (z_1 \cdot z_2) = \text{Ln } z_1 + \text{Ln } z_2$ (если учесть, что равные значения показательная функция принимает лишь в точках, отличающихся между собой на число, кратное $2\pi i$).

Заметим, что запись $\text{Ln } z_1 + \text{Ln } z_2$ следует понимать так: каждое значение $\text{Ln } z_1$ складывается с каждым значением $\text{Ln } z_2$ и полученное таким образом множество комплексных чисел совпадает с множеством значений $\text{Ln } (z_1 \cdot z_2)$.

Степенная функция с комплексным показателем. Степенная функция $w = z^\alpha$ (где α — произвольное комплексное число и $z \neq 0$) определяется, равенством $z^\alpha = e^{\alpha \text{Ln } z}$.

Так как $\text{Ln } z$ многозначная функция, то то же самое, вообще говоря, можно сказать и относительно функции z^α . Действительно, записав $\text{Ln } z = \ln z + 2k\pi i$, получим

$$z^\alpha = e^{\alpha \text{Ln } z} = e^{\alpha \ln z + 2k\pi i \cdot \alpha} = e^{\alpha \ln z} \cdot e^{2k\pi i \cdot \alpha},$$

и если обозначить одно из значений z^α , при фиксированном z , через w_0 , $w_0 = e^{\alpha \ln z}$, то все значения z^α при этом значении аргумента определяются равенством

$$z^\alpha = w_0 e^{2k\pi i \cdot \alpha}.$$

Отсюда заключаем:

1) Если α — целое число, то $e^{2k\pi i \cdot \alpha} = 1$ и функция z^α имеет единственное значение. В частности, при $\alpha=0$ $z^\alpha = 1$.

2) Если α — дробное рациональное число, $\alpha = \frac{m}{n}$ (причем $\frac{m}{n}$ — несократимая дробь), то учитывая периодичность показательной функции, получим, что $e^{2k\pi i \alpha} = e^{\frac{2k\pi i m}{n}}$ имеет только n различных значений, соответствующих, например, значениям $k = 0, 1, 2, \dots, n-1$. В частности, при $m=1$, получим $z^n = \sqrt[n]{z}$. Эта функция имеет n различных значений. (Этот результат был получен также в гл. V, § 7.)

3) Если же α — иррациональное или комплексное число, то можно показать, что функция z^α при фиксированном z имеет бесчисленное множество значений.

Пример 1. $1^i = e^{i \operatorname{Ln} 1} = e^{i \cdot 2k\pi} = e^{-2k\pi} \quad (k=0, \pm 1, \pm 2, \dots)$, в частности, при $k=0$ $1^i = 1$.

Пример 2. $i^i = e^{i \operatorname{Ln} i} = e^{i \cdot i \left(\frac{\pi}{2} + 2k\pi \right)} = e^{-\left(\frac{\pi}{2} + 2k\pi \right)} \quad (k=0, \pm 1, \pm 2, \dots)$.

Обратные тригонометрические функции комплексного аргумента. Обратные тригонометрические функции обозначаются

$$\operatorname{Arcsin} z, \operatorname{Arccos} z, \operatorname{Arctg} z$$

и определяются следующим образом: $w = \operatorname{Arcsin} z$, если $\sin w = z$; $w = \operatorname{Arccos} z$, если $\cos w = z$ и $w = \operatorname{Arctg} z$, если $\operatorname{tg} w = z$.

Как известно, тригонометрические функции выражаются через показательную функцию, поэтому естественно ожидать, что функции, обратные тригонометрическим, будут выражаться через логарифм. Действительно, по определению $w = \operatorname{Arcsin} z$, если $z = \sin w$ или

$$z = \frac{e^{iw} - e^{-iw}}{2i},$$

откуда получаем:

$$e^{2iw} - 2ize^{iw} - 1 = 0.$$

Решая это квадратное уравнение относительно e^{iw} , имеем:

$$\begin{aligned} e^{iw} &= iz + \sqrt{1 - z^2}, \\ w &= \frac{1}{i} \operatorname{Ln} (iz + \sqrt{1 - z^2}) \end{aligned}$$

или, что то же,

$$\operatorname{Arcsin} z = \frac{1}{i} \operatorname{Ln} (iz + \sqrt{1 - z^2}).$$

Все значения функции $\operatorname{Arcsin} z$ определяются этим равенством.

Так как выражение, записанное под знаком логарифма, отлично от нуля, то функция $\operatorname{Arcsin} z$ определена при всех значениях z . $\operatorname{Arcsin} z$ — бесконечнозначная функция, что объясняется двухзначностью корня и бесконечнозначностью логарифма.

Аналогично получаются выражения через логарифм для функций: $\operatorname{Arccos} z$ и $\operatorname{Arctg} z$. Действительно, $w = \operatorname{Arccos} z$, если $z = \cos w = \frac{e^{iw} + e^{-iw}}{2}$, откуда $e^{2iw} - 2ze^{iw} + 1 = 0$. Решая уравнение, найдем $e^{iw} = z + \sqrt{z^2 - 1}$ и, следовательно,

$$w = \frac{1}{i} \operatorname{Ln} (z + \sqrt{z^2 - 1}).$$

Таким образом, все значения $\operatorname{Arccos} z$ определяются равенством

$$\operatorname{Arccos} z = \frac{1}{i} \operatorname{Ln} (z + \sqrt{z^2 - 1}).$$

Это — бесконечнозначная функция, определенная при всех значениях z (так как выражение, записанное под знаком логарифма, в нуль не обращается). Бесконечнозначность этой функции объясняется двузначностью корня и бесконечнозначностью логарифма.

$$\text{И, наконец, } w = \operatorname{Arctg} z, \text{ если } z = \operatorname{tg} w = \frac{e^{i\omega} - e^{-i\omega}}{i(e^{i\omega} + e^{-i\omega})}$$

или

$$iz = \frac{e^{2i\omega} - 1}{e^{2i\omega} + 1},$$

откуда

$$e^{2i\omega} = \frac{1 + iz}{1 - iz} = \frac{i - z}{i + z}.$$

$$\text{И окончательно } w = \operatorname{Arctg} z = \frac{1}{2i} \operatorname{Ln} \frac{i - z}{i + z}.$$

Это — бесконечнозначная функция (за счет логарифма), определенная при всех значениях z , кроме $z = \pm i$.

Пример. Найти нули $\sin z$, т. е. определить z , при которых $\sin z = 0$.

Решение. $\sin z = 0$ при $z = \operatorname{Arcsin} 0 = \frac{1}{i} \operatorname{Ln} (\sqrt{1})$, откуда получаем две системы значений, которые обозначим $(z)_1$ и $(z)_2$:

$$(z)_1 = \frac{1}{i} \operatorname{Ln} 1 = \frac{1}{i} 2k\pi i = 2k\pi,$$

$$(z)_2 = \frac{1}{i} \operatorname{Ln} (-1) = \frac{1}{i} (\pi + 2k\pi) i = (2k + 1)\pi$$

или, что то же, $\sin z = 0$ при $z = k\pi$ ($k = 0, \pm 1, \pm 2, \dots$).

Обратные гиперболические функции. Обратные гиперболические функции обозначаются¹⁾ $\operatorname{Arsh} z$ (арсинус гиперболический z), $\operatorname{Arch} z$ (аркосинус гиперболический z), $\operatorname{Arth} z$ (артангенс гиперболический z) и определяются следующим образом: $w = \operatorname{Arsh} z$, если $\operatorname{sh} w = z$; $w = \operatorname{Arch} z$, если $\operatorname{ch} w = z$ и $w = \operatorname{Arth} z$, если $\operatorname{th} w = z$.

Найдем выражения этих функций через логарифм.

По определению $w = \operatorname{Arsh} z$, если $z = \operatorname{sh} w = \frac{e^w - e^{-w}}{2}$ или $e^{2w} - 2ze^w - 1 = 0$, откуда $e^w = z + \sqrt{z^2 + 1}$ и, следовательно, $w = \operatorname{Arsh} z = \operatorname{Ln} (z + \sqrt{z^2 + 1})$.

Таким образом, $\operatorname{Arsh} z$ — бесконечнозначная функция, определенная при всех значениях z . Она бесконечнозначна за счет двузначности корня и бесконечнозначности логарифма.

По определению $w = \operatorname{Arch} z$, если $z = \operatorname{ch} w = \frac{e^w + e^{-w}}{2}$, и, проводя аналогичные рассуждения, получим:

$$\operatorname{Arch} z = \operatorname{Ln} (z + \sqrt{z^2 - 1}).$$

И, наконец, если $w = \operatorname{Arth} z$, то $z = \operatorname{th} w = \frac{e^w - e^{-w}}{e^w + e^{-w}}$, и, следовательно, можно получить: $\operatorname{Arth} z = \frac{1}{2} \operatorname{Ln} \frac{1 + z}{1 - z}$.

Функция $\operatorname{Arth} z$ бесконечнозначная (за счет логарифма), определенная при всех значениях $z \neq \pm 1$.

¹⁾ Ar — от Area — означает (лат.) площадь.

Раздел 5

РЯДЫ И ИНТЕГРАЛ ФУРЬЕ

§ 16. Основные сведения. Тригонометрические ряды Фурье для функций с периодом 2π

Функция $f(x)$, определенная при всех значениях x , называется *периодической*, если существует такое число T ($T \neq 0$), что при любом значении x выполняется равенство

$$f(x+T)=f(x). \quad (1)$$

Число T в этом случае называется *периодом функции*. Как следует из определения периодической функции, для построения графика такой функции достаточно построить ее график на отрезке $[0, T]$ и периодически продолжить его на всю числовую ось (рис. 440). (Прибавление числа к аргументу функции означает сдвиг графика этой функции вдоль оси OX .)

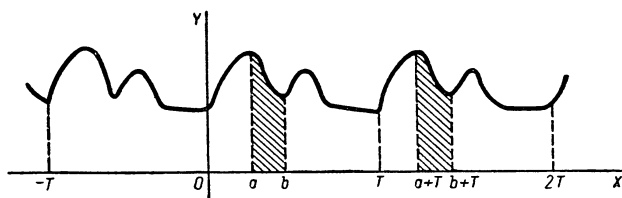


Рис. 440.

Из определения также следует, что если T — период функции, то ее периодом также будет nT , где n — любое целое число. Обычно за период T принимают наименьшее положительное число, удовлетворяющее равенству (1).

Отметим следующие свойства периодических функций:

1) Сумма, разность, произведение и частное периодических функций периода T есть периодические функции периода T .

2) Если функция $f(x)$ имеет период T , то функция $f(ax)$ имеет период $\frac{T}{a}$.

В самом деле, для любого x :

$$f\left[a\left(x+\frac{T}{a}\right)\right]=f(ax+T)=f(ax)$$

(умножение аргумента на число означает сжатие или растяжение графика этой функции вдоль оси OX).

Например, $\cos 3x$ имеет период $T=\frac{2\pi}{3}$, периодом $\sin nx$ будет $T=\frac{2\pi}{n}$.

3) Если $f(x)$ периодическая функция периода T , то равны любые два интеграла от этой функции, взятые по промежутку длины T (при этом предполагается, что эти интегралы существуют). Итак, докажем, что при любых a и b справедливо равенство

$$\int_a^{a+T} f(x) dx = \int_b^{b+T} f(x) dx$$

(геометрически для неотрицательных функций имеем равенство площадей, рис. 440). Действительно,

$$\int_a^{a+T} f(x) dx = \int_a^b f(x) dx + \int_b^{b+T} f(x) dx + \int_{b+T}^{a+T} f(x) dx.$$

Преобразуем теперь интеграл

$$\int_{b+T}^{a+T} f(x) dx,$$

полагая $x = u + T$:

$$\int_{b+T}^{a+T} f(x) dx = \int_b^a f(u+T) du = \int_b^a f(u) du$$

(при этом мы воспользовались периодичностью функции $f(u)$). И окончательно получаем:

$$\int_a^{a+T} f(x) dx = \int_a^b f(x) dx + \int_b^{b+T} f(x) dx + \int_b^a f(x) dx = \int_b^{b+T} f(x) dx.$$

Тригонометрический ряд. Ряд Фурье. Тригонометрическим рядом называется ряд вида:

$$\frac{a_0}{2} + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \\ + \dots + (a_n \cos nx + b_n \sin nx) + \dots$$

или, короче,

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

где $a_0, a_1, b_1, a_2, b_2, \dots, a_n, b_n, \dots$ — действительные числа, называемые *коэффициентами ряда*¹⁾. Каждое слагаемое $a_n \cos nx + b_n \sin nx$ ($n=1, 2, \dots$) тригонометрического ряда является периодической функцией периода 2π (так как $\frac{a_0}{2}$ — имеет любой период, а период $a_n \cos nx + b_n \sin nx$ равен $\frac{2\pi}{n}$, а значит, и 2π). Каждое слагаемое $a_n \cos nx + b_n \sin nx$ ($n=1, 2, \dots$) является аналитическим выражением простого гармонического колебания $A \sin(nx + \varphi)$, где $A \sin n\varphi = a_n$, $A \cos n\varphi = b_n$, A — амплитуда, φ — начальная фаза. Учитывая сказанное, получаем: если тригонометрический ряд сходится на отрезке длины периода 2π , то он сходится на всей числовой оси и его сумма $f(x)$ является периодической функцией периода 2π . А также получаем, что в этом случае периодическая функция $f(x)$ записана в виде суммы бесконечного ряда простых гармоник или, что то же, разложена в ряд простых гармоник.

Пусть тригонометрический ряд равномерно сходится на отрезке $[-\pi, \pi]$ (следовательно, и на любом отрезке) и его сумма равна $f(x)$. Ставится задача: определить коэффициенты этого ряда.

Для решения поставленной задачи воспользуемся следующими равенствами:

$$\int_{-\pi}^{\pi} \cos mx \cos nx dx = \begin{cases} 0, & m \neq n \\ \pi, & m = n, \quad m, n = 0, 1, 2, \dots, \end{cases}$$

$$\int_{-\pi}^{\pi} \sin mx \sin nx dx = \begin{cases} 0, & m \neq n \\ \pi, & m = n, \quad m, n = 0, 1, 2, \dots, \end{cases}$$

$$\int_{-\pi}^{\pi} \cos mx \sin nx dx = 0, \quad m = 0, 1, 2, \dots, \quad n = 1, 2, \dots$$

¹⁾ Как будет видно дальше, свободный член удобно записывать в виде $\frac{a_0}{2}$.

Нетрудно убедиться в справедливости записанных равенств. В самом деле,

$$\cos mx \cos nx = \frac{1}{2} [\cos (m+n)x + \cos (m-n)x],$$

$$\sin mx \sin nx = \frac{1}{2} [\cos (m-n)x - \cos (m+n)x],$$

$$\cos mx \sin nx = \frac{1}{2} [\sin (n+m)x - \sin (n-m)x]$$

и, таким образом, все интегралы приводятся к интегралам вида:

$$\int_{-\pi}^{\pi} \cos kx \, dx \quad \text{и} \quad \int_{-\pi}^{\pi} \sin kx \, dx, \quad \text{где} \quad k=0, 1, 2, 3, \dots,$$

$$\int_{-\pi}^{\pi} \cos kx \, dx = \begin{cases} x \Big|_{-\pi}^{\pi} = 2\pi, & \text{при } k=0, \\ \frac{\sin kx}{k} \Big|_{-\pi}^{\pi} = 0, & \text{при } k \neq 0, \end{cases}$$

$$\int_{-\pi}^{\pi} \sin kx \, dx = \begin{cases} 0, & \text{при } k=0, \\ -\frac{\cos kx}{k} \Big|_{-\pi}^{\pi} = 0, & \text{при } k \neq 0. \end{cases}$$

А также воспользуемся следующими свойствами:

1) Как известно, сумма равномерно сходящегося на некотором отрезке $[a, b]$ ряда непрерывных функций является непрерывной функцией на этом отрезке. Учитывая это, получим, что сумма равномерно сходящегося на отрезке $[-\pi, \pi]$ тригонометрического ряда — непрерывная функция на всей числовой оси.

2) Равномерная сходимость ряда $\sum_{n=1}^{\infty} u_n(x)$ на отрезке $[a, b]$ не нарушится, если все члены ряда умножить на функцию $\varphi(x)$, непрерывную на этом отрезке.

Действительно, применяя к ряду $\sum_{n=1}^{\infty} u_n(x)$ критерий Коши равномерной сходимости, получим: для любого числа $\varepsilon > 0$ найдется такой номер $N = N(\varepsilon)$, что при $n > N$ и любом целом $p > 0$ неравенство

$$|u_{n+1}(x) + u_{n+2}(x) + \dots + u_{n+p}(x)| < \varepsilon$$

выполняется для всех x отрезка $[a, b]$.

По условию $\varphi(x)$ — непрерывная функция на отрезке $[a, b]$, следовательно, она ограничена на этом отрезке, т. е. существует такое число $M > 0$, что для всех x отрезка $[a, b]$ выполняется неравенство $|\varphi(x)| \leq M$.

Тогда для ряда $\sum_{n=1}^{\infty} \varphi(x) u_n(x)$, полученного из данного умножением на функцию $\varphi(x)$, непрерывную на отрезке $[a, b]$, имеет место неравенство $|\varphi(x) u_{n+1}(x) + \varphi(x) u_{n+2}(x) + \dots + \varphi(x) u_{n+p}(x)| < M\varepsilon$, справедливое для всех x отрезка $[a, b]$ при $n > N$ и любом целом $p > 0$. А следовательно, по критерию Коши ряд $\sum_{n=1}^{\infty} \varphi(x) u_n(x)$ равномерно сходится на отрезке $[a, b]$, что и требовалось доказать.

В частности, равномерная сходимость на отрезке $[-\pi, \pi]$ данного тригонометрического ряда не нарушится, если ряд умножить на $\cos nx$ или $\sin nx$.

Теперь перейдем к решению поставленной задачи. По условию

$$f(x) = \frac{a_0}{2} + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \dots + (a_n \cos nx + b_n \sin nx) + \dots, \quad (2)$$

при этом ряд равномерно сходится на отрезке $[-\pi, \pi]$. Как было получено, $f(x)$ — непрерывная функция на отрезке $[-\pi, \pi]$, а следовательно, интеграл $\int_{-\pi}^{\pi} f(x) dx$ существует.

В результате почленного интегрирования равномерно сходящегося ряда (2) получаем:

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) dx = & \frac{a_0}{2} \int_{-\pi}^{\pi} dx + \left(a_1 \int_{-\pi}^{\pi} \cos x dx + b_1 \int_{-\pi}^{\pi} \sin x dx \right) + \\ & + \left(a_2 \int_{-\pi}^{\pi} \cos 2x dx + b_2 \int_{-\pi}^{\pi} \sin 2x dx \right) + \\ & + \dots + \left(a_n \int_{-\pi}^{\pi} \cos nx dx + b_n \int_{-\pi}^{\pi} \sin nx dx \right) + \dots = \pi a_0 \end{aligned}$$

откуда

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx.$$

Умножая равенство (2) на $\cos nx$ и интегрируя его в пределах от $-\pi$ до π , получаем:

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) \cos nx dx = & \frac{a_0}{2} \int_{-\pi}^{\pi} \cos nx dx + \left(a_1 \int_{-\pi}^{\pi} \cos x \cos nx dx + \right. \\ & + b_1 \int_{-\pi}^{\pi} \sin x \cos nx dx \left. \right) + \left(a_2 \int_{-\pi}^{\pi} \cos 2x \cos nx dx + \right. \\ & + b_2 \int_{-\pi}^{\pi} \sin 2x \cos nx dx \left. \right) + \dots + \left(a_n \int_{-\pi}^{\pi} \cos^2 nx dx + \right. \\ & + b_n \int_{-\pi}^{\pi} \sin nx \cos nx dx \left. \right) + \dots = \pi a_n, \end{aligned}$$

тогда

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \quad (n = 1, 2, \dots).$$

Аналогично, умножая равенство (2) на $\sin nx$ и интегрируя его в пределах от $-\pi$ до π , имеем:

$$\int_{-\pi}^{\pi} f(x) \sin nx dx = \pi b_n,$$

следовательно,

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx.$$

Таким образом, получены следующие выражения для коэффициентов ряда:

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx, \\ a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \quad (n = 1, 2, \dots), \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \quad (n = 1, 2, \dots). \end{aligned}$$

Или, объединяя первые два равенства, запишем:

$$\left. \begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \quad (n=0, 1, 2, \dots), \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \quad (n=1, 2, \dots). \end{aligned} \right\} \quad (3)$$

Заметим, что в равенствах (3) пределы интегрирования можно взять от 0 до 2π (на основании свойства периодических функций) или от a до $a+2\pi$, где a — произвольное число.

Пусть $f(x)$ — любая периодическая функция периода 2π , непрерывная на отрезке $[-\pi, \pi]$ или имеющая на этом отрезке конечное число точек разрыва первого ряда.

Тогда для функции $f(x)$ коэффициенты, определяемые равенствами (3), существуют и называются *коэффициентами Фурье этой функции*¹⁾.

Рядом Фурье такой функции $f(x)$ называется тригонометрический ряд, коэффициенты которого определяются равенствами (3). Из приведенных выше рассуждений непосредственно получаем: во-первых, если $f(x)$ является суммой равномерно сходящегося на отрезке $[-\pi, \pi]$ тригонометрического ряда, то этот ряд является рядом Фурье функции $f(x)$; во-вторых, не существует двух тригонометрических рядов, равномерно сходящихся на отрезке $[-\pi, \pi]$ к одной и той же функции [см. равенства (3)]. Или, что то же, если функция $f(x)$ разложена в равномерно сходящийся на отрезке $[-\pi, \pi]$ тригонометрический ряд, то такое разложение единственное.

Пусть $f(x)$ — любая периодическая функция периода 2π , непрерывная на отрезке $[-\pi, \pi]$ или имеющая на этом отрезке конечное число точек разрыва первого рода. Для такой функции $f(x)$ можно записать соответствующий ей ряд Фурье:

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

коэффициенты которого определяются из равенств (3). Однако, как и в случае степенных рядов, нельзя утверждать, что ряд Фурье функции $f(x)$ сходится к этой функции.

Этот ряд может для некоторых значений x оказаться расходящимся, но если даже он и сходится, то это еще не означает, что его сумма равна $f(x)$. Если ряд Фурье функции $f(x)$ сходится к ней во всех ее точках непрерывности, то знак соответствия $\sim$ заменяется знаком равенства, и в этом случае говорят, что *функция $f(x)$ разлагается в ряд Фурье*.

Можно доказать, что при некоторых условиях, налагаемых на функцию $f(x)$, она разлагается в ряд Фурье. Опуская доказательства, сформулируем только некоторые достаточные признаки разложимости функции в ряд Фурье.

Достаточные признаки разложимости функции в ряд Фурье. Прежде всего напомним, что точку x_0 разрыва функции $f(x)$ называют точкой разрыва первого рода, если существуют конечные пределы справа и слева функции $f(x)$ в точке x_0 :

$$\begin{aligned} \lim_{x \rightarrow x_0+0} f(x) &= f(x_0+0) \text{ — предел справа,} \\ \lim_{x \rightarrow x_0-0} f(x) &= f(x_0-0) \text{ — предел слева.} \end{aligned}$$

¹⁾ Фурье (1768—1830) — французский математик.

Теорема 1 (Дирихле). Если функция $f(x)$ имеет период 2π и на отрезке $[-\pi, \pi]$ непрерывна или имеет конечное число точек разрыва первого рода, и отрезок $[-\pi, \pi]$ можно разбить на конечное число отрезков так, что внутри каждого из них $f(x)$ монотонна, то ряд Фурье функции $f(x)$ сходится при всех значениях x , причем в точках непрерывности функции $f(x)$ его сумма равна $f(x)$, а в точках разрыва функции $f(x)$ его сумма равна $\frac{f(x-0) + f(x+0)}{2}$, т. е. среднему арифметическому предельных значений слева и справа. Кроме того, ряд Фурье функции $f(x)$ сходится равномерно на любом отрезке, который вместе со своими концами принадлежит интервалу непрерывности функции $f(x)$.

Функция, удовлетворяющая условиям теоремы Дирихле (кроме условия периодичности), называется кусочно-монотонной на отрезке $[-\pi, \pi]$.

Теорема 2. Если функция $f(x)$ имеет период 2π , кроме того, $f(x)$ и ее производная $f'(x)$ — непрерывные функции на отрезке $[-\pi, \pi]$ или имеют конечное число точек разрыва первого рода на этом отрезке, то ряд Фурье функции $f(x)$ сходится при всех значениях x , причем в точках непрерывности функции $f(x)$ его сумма равна $f(x)$, а в точках разрыва функции $f(x)$ она равна $\frac{f(x-0) + f(x+0)}{2}$. И, наконец, ряд Фурье функции $f(x)$ сходится равномерно на любом отрезке, который вместе со своими концами принадлежит интервалу непрерывности функции $f(x)$.

Функция $f(x)$, удовлетворяющая условиям этой теоремы, кроме условия периодичности, называется кусочно-гладкой на отрезке $[-\pi, \pi]$ ¹⁾.

Следует иметь в виду, что существуют и другие достаточные признаки разложимости функции в ряд Фурье, но для решения практических задач обычно пользуются приведенными здесь признаками. Заметим, что сформулированные признаки между собой, вообще говоря, не сравнимы, так как существуют функции, удовлетворяющие теореме 1 и не удовлетворяющие теореме 2, и наоборот.

Например, функция $f(x)$, определенная на отрезке $[-\pi, \pi]$ равенством

$$f(x) = \begin{cases} x^4 \sin \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0 \end{cases}$$

и подчиненная условию $f(x+2\pi) = f(x)$, удовлетворяет условиям теоремы 2 и не удовлетворяет условиям теоремы 1. С другой стороны, существуют непрерывные на отрезке $[-\pi, \pi]$ монотонные функции, не имеющие производных в бесчисленном множестве точек отрезка $[-\pi, \pi]$, т. е. не удовлетворяющие условиям теоремы 2, но удовлетворяющие условиям теоремы 1, если подчинить их еще условию периодичности:

$$f(x+2\pi) = f(x).$$

Рассмотрим теперь вопрос о разложении в ряд Фурье *непериодической функции*, заданной на отрезке (интервале) длиной 2π или меньше 2π .

Пусть функция $f(x)$ задана на отрезке $[a, a+2\pi]$ и является кусочно-гладкой или кусочно-монотонной на этом отрезке и, кроме того, $f(a) = f(a+2\pi)$. Периодически продолжая эту функцию на всю числовую ось, т. е. подчиняем ее условию $f(x+2\pi) = f(x)$. Тогда полученную периодическую функцию можно разложить в ряд Фурье, который будет

¹⁾ Если вместо отрезка $[-\pi, \pi]$ рассмотреть отрезок $[a, b]$, то получим функции соответственно кусочно-гладкие, кусочно-монотонные на отрезке $[a, b]$.

сходиться к заданной функции $f(x)$ в каждой ее точке непрерывности, принадлежащей отрезку $[a, a + 2\pi]$.

Пусть теперь функция $f(x)$, заданная на интервале $(a, a + 2\pi)$, удовлетворяет следующему условию: существует на отрезке $[a, a + 2\pi]$ кусочно-гладкая или кусочно-монотонная функция, совпадающая с $f(x)$ на интервале $(a, a + 2\pi)$. При этом условии функция, которая является периодическим продолжением функции $f(x)$, разлагается в ряд Фурье, сходящийся к $f(x)$ в каждой ее точке непрерывности, принадлежащей интервалу $(a, a + 2\pi)$.

Таким образом, задача разложения в ряд Фурье функции, заданной на отрезке (интервале) длины 2π , фактически ничем не отличается от задачи разложения в ряд Фурье периодической функции.

Если такая же задача решается по отношению к функции $f(x)$, заданной на отрезке $[a, a + \lambda]$ [интервале $(a, a + \lambda)$]¹⁾, длина которого $\lambda < 2\pi$, то она приводится к предшествующей задаче, если только продолжить функцию $f(x)$ на интервал $(a + \lambda, a + 2\pi)$ произвольным образом²⁾, но так, чтобы условия разложимости функции в ряд Фурье сохранялись и выполнялось равенство $f(a + 2\pi) = f(a)$. (В случае функ-

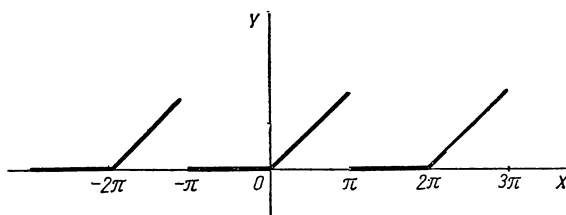


Рис. 441.

ции, заданной на интервале, естественно, последнее условие отпадает). Тогда полученная таким образом функция разлагается в ряд Фурье на отрезке $[a, a + \lambda]$ [интервале $(a, a + \lambda)$], так как сумма этого ряда на отрезке $[a, a + \lambda]$ [интервале $(a, a + \lambda)$] равна заданной функции $f(x)$ в каждой ее точке непрерывности.

Следует иметь в виду, что при $\lambda < 2\pi$ продолжение заданной функции $f(x)$ на отрезок (интервал) длиной 2π возможно бесчисленным множеством способов, поэтому заданная функция в области ее определения бесчисленным множеством способов может быть разложена в ряд Фурье. Ясно, что суммы этих рядов будут различными функциями на отрезке $[a, a + 2\pi]$, но они будут совпадать с заданной функцией $f(x)$ на отрезке $[a, a + \lambda]$ [интервале $(a, a + \lambda)$], кроме, возможно, конечного числа точек разрыва $f(x)$.

Пример. Разложить в ряд Фурье функцию

$$f(x) = \begin{cases} x, & 0 \leq x < \pi, \\ 0, & -\pi < x < 0, \end{cases}$$

удовлетворяющую условию $f(x + 2\pi) = f(x)$ (рис. 441).

Решение. Функция $f(x)$ удовлетворяет условиям разложимости в ряд Фурье, поэтому можно записать:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

¹⁾ При соответствующих условиях на функцию $f(x)$.

²⁾ Или, что то же, доопределить функцию $f(x)$ на интервале $(a + \lambda, a + 2\pi)$.

при этом

$$\begin{aligned}
 a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{1}{\pi} \int_0^{\pi} x dx = \frac{1}{\pi} \left[\frac{x^2}{2} \right]_0^{\pi} = \frac{\pi}{2}, \\
 a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx = \frac{1}{\pi} \int_0^{\pi} x \cos nx dx = \\
 &= \frac{1}{\pi} \left[\frac{x \sin nx}{n} \right]_0^{\pi} - \frac{1}{n} \int_0^{\pi} \sin nx dx = \frac{\cos nx}{n^2 \pi} \Big|_0^{\pi} = \\
 &= \frac{\cos n\pi - 1}{n^2 \pi} = \frac{(-1)^n - 1}{n^2 \pi} = \begin{cases} 0 & \text{при } n \text{ четном,} \\ -\frac{2}{n^2 \pi} & \text{при } n \text{ нечетном,} \end{cases} \\
 a_n &= \begin{cases} 0 & \text{при } n \text{ четном,} \\ -\frac{2}{n^2 \pi} & \text{при } n \text{ нечетном,} \end{cases} \\
 b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = \frac{1}{\pi} \int_0^{\pi} x \sin nx dx = \\
 &= \frac{1}{\pi} \left[-\frac{x \cos nx}{n} \right]_0^{\pi} + \frac{1}{n} \int_0^{\pi} \cos nx dx = \frac{1}{\pi} \left[-\frac{x \cos nx}{n} + \frac{\sin nx}{n^2} \right]_0^{\pi} = \\
 &= -\frac{\cos n\pi}{n} = \frac{(-1)^{n+1}}{n}; \quad b_n = \frac{(-1)^{n+1}}{n} \quad (n = 1, 2, \dots).
 \end{aligned}$$

И, следовательно,

$$\begin{aligned}
 f(x) &= \frac{\pi}{4} + \left(-\frac{2}{\pi} \cos x + \sin x \right) - \frac{\sin 2x}{2} + \left(-\frac{2}{9\pi} \cos 3x + \frac{1}{3} \sin 3x \right) - \\
 &\quad - \frac{\sin 4x}{4} + \left(-\frac{2}{25\pi} \cos 5x + \frac{1}{5} \sin 5x \right) + \dots
 \end{aligned} \quad (1'')$$

Полагая в этом равенстве $x=0$, получим:

$$0 = \frac{\pi}{4} - \frac{2}{\pi} \left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots \right),$$

откуда

$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = \frac{\pi^2}{8}.$$

Используя этот результат, нетрудно убедиться в том, что сумма ряда, записанного в правой части равенства (1''), при $x=\pi$ (а следовательно, и при $x=(2k+1)\pi$, k — любое целое число) равна $\frac{\pi}{2}$, т. е. равна $\frac{f(\pi-0) + f(\pi+0)}{2}$.

Ряды Фурье для четных и нечетных функций с периодом 2π . Прежде всего отметим, что в дальнейшем нам придется пользоваться следующим свойством интеграла по симметричному относительно $x=0$ промежутку:

$$\int_{-a}^a f(x) dx = \begin{cases} 0, & \text{если } f(x) \text{ — нечетная функция,} \\ 2 \int_0^a f(x) dx, & \text{если } f(x) \text{ — четная функция.} \end{cases}$$

В самом деле,

$$\int_{-a}^a f(x) dx = \int_{-a}^0 f(x) dx + \int_0^a f(x) dx.$$

Сделаем в первом интеграле подстановку $x = -t$, тогда

$$\int_{-a}^0 f(x) dx = - \int_a^0 f(-t) dt = \int_0^a f(-t) dt = \int_0^a f(-x) dx,$$

при этом

$$\int_{-a}^a f(x) dx = \int_0^a [f(x) + f(-x)] dx.$$

Если $f(x)$ — четная функция, то $f(-x) = f(x)$ и, следовательно,

$$\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx.$$

Если же $f(x)$ — нечетная функция, то $f(-x) = -f(x)$ и, следовательно,

$$\int_{-a}^a f(x) dx = 0.$$

Заметим, что произведение двух четных или двух нечетных функций — четная функция, а произведение четной функции на нечетную — нечетная функция.

Пусть теперь $f(x)$ — четная периодическая функция с периодом 2π , удовлетворяющая условиям разложимости в ряд Фурье. Тогда, используя только что полученное свойство интегралов, получим:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos nx dx \quad (n=0, 1, 2, \dots),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = 0 \quad (n=1, 2, \dots).$$

Таким образом, ряд Фурье для четной функции содержит только четные функции — косинусы и записывается так:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx, \quad (3)$$

при этом

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(x) \cos nx dx \quad (n=0, 1, 2, \dots). \quad (4)$$

Рассуждая аналогично, получаем, что если $f(x)$ — нечетная периодическая ($T=2\pi$) функция, удовлетворяющая условиям разложимости в ряд Фурье, то

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx = 0 \quad (n=0, 1, 2, \dots),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx dx \quad (n=1, 2, \dots).$$

Следовательно, ряд Фурье нечетной функции содержит только нечетные функции — синусы и записывается следующим образом:

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx, \quad (5)$$

при этом

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx dx \quad (n=1, 2, \dots).$$

Заметим, что если функция $f(x)$ задана на отрезке $[0, \pi]$ [интервале $(0, \pi)$] и удовлетворяет в рассматриваемом промежутке условиям разложимости в ряд Фурье (см. п. 2), то в этом промежутке ее можно бесчисленным множеством способов разложить в ряд Фурье. В частности, функцию $f(x)$ можно разложить в ряд по косинусам или по синусам, для этого нужно только продолжить функцию $f(x)$ на отрезок $[-\pi, 0]$ [интервал $(-\pi, 0)$] соответственно четным или нечетным образом, т. е. подчинить, соответственно условию

$$f(-x) = f(x) \quad \text{или} \quad f(-x) = -f(x).$$

Пример 1. Разложить в ряд Фурье периодическую ($T=2\pi$) функцию

$$f(x) = \begin{cases} -1 & \text{на интервале } (-\pi, 0), \\ 1 & \text{на интервале } (0, \pi) \end{cases}$$

(рис. 442).

Пусть в точках $x = k\pi$ (k — целое число) значения функции равны 0, $f(k\pi) = 0$. Сумма ряда Фурье для функции $f(x)$ в этих точках также будет равна 0, т. е. будет

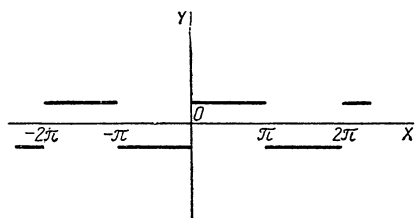


Рис. 442.

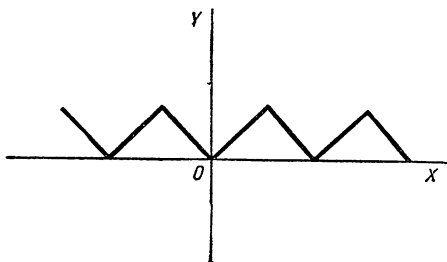


Рис. 443.

равна среднему арифметическому пределов слева и справа. Так как заданная нечетная функция $f(x)$ удовлетворяет условиям разложимости в ряд Фурье, то

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx,$$

при этом

$$\begin{aligned} b_n &= \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx \, dx = \frac{2}{\pi} \int_0^{\pi} \sin nx \, dx = \\ &= -\frac{2}{n\pi} \cos nx \Big|_0^{\pi} = \frac{2}{n\pi} (1 - \cos n\pi) \end{aligned}$$

или, что то же,

$$b_n = \frac{2}{n\pi} [1 - (-1)^n] = \begin{cases} 0 & \text{при } n \text{ четном,} \\ \frac{4}{n\pi} & \text{при } n \text{ нечетном,} \end{cases}$$

и ряд Фурье для функции $f(x)$ запишется так:

$$f(x) = \frac{4}{\pi} \left(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \dots \right). \quad (2')$$

При $x = k\pi$ ($k = 0, \pm 1, \pm 2, \dots$) сумма ряда, записанного в правой части равенства, как уже отмечалось выше, равна 0. Полагая $x = \frac{\pi}{2}$, получим:

$$1 = \frac{4}{\pi} \left(1 - \frac{1}{3} + \frac{1}{5} - \dots + \frac{(-1)^{n+1}}{(2n-1)} + \dots \right),$$

откуда

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)} = \frac{\pi}{4}.$$

Рекомендуется построить графики частных сумм $S_1(x)$, $S_2(x)$, $S_3(x)$ ряда (2') и посмотреть, как они приближаются к графику функции $f(x)$.

Пример 2. Разложить в ряд Фурье по косинусам функцию $f(x) = x$ на отрезке $[0, \pi]$.

Решение. Из условия задачи непосредственно следует, что нужно сделать четное продолжение заданной функции на отрезок $[-\pi, 0]$. Далее периодически продолжить полученную таким образом функцию на всю числовую ось (рис. 443) и затем разложить эту функцию в ряд Фурье. Обозначим периодическую функцию, изображенную на рис. 443, через $\varphi(x)$. Четная функция $\varphi(x)$ удовлетворяет условиям разложимости в ряд Фурье, следовательно, по формулам (3), (4) имеем:

$$\varphi(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx,$$

при этом

$$a_0 = \frac{2}{\pi} \int_0^{\pi} x \, dx = \frac{2}{\pi} \left. \frac{x^2}{2} \right|_0^{\pi} = \pi,$$

$$\begin{aligned} a_n &= \frac{2}{\pi} \int_0^{\pi} x \cos nx \, dx = \frac{2}{\pi} \left[\frac{x \sin nx}{n} + \frac{\cos nx}{n^2} \right]_0^{\pi} = \frac{2}{n^2 \pi} (\cos n\pi - 1) = \\ &= \frac{2}{n^2 \pi} [(-1)^n - 1] = \begin{cases} 0 & \text{при } n \text{ четном,} \\ -\frac{4}{n^2 \pi} & \text{при } n \text{ нечетном,} \end{cases} \end{aligned}$$

и окончательно получаем:

$$\varphi(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cos (2n-1)x}{(2n-1)^2}.$$

На отрезке $[0, \pi]$ функция $\varphi(x)$ совпадает с функцией $f(x)$ и, значит,

$$x = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cos (2n-1)x}{(2n-1)^2}, \quad 0 \leq x \leq \pi.$$

Следует иметь в виду, что нельзя забыть «приписать» условие $0 \leq x \leq \pi$, иначе полученное равенство будет неверно, хотя бы потому, что сумма ряда — периодическая функция, а x — непериодическая функция.

Заметим еще, что функция $\varphi(x) = |x|$ при $-\pi \leq x \leq \pi$ и, значит, мы решили также задачу разложения в ряд Фурье периодической функции с периодом 2π , равной $|x|$ на отрезке $[-\pi, \pi]$.

§ 17. Тригонометрические ряды Фурье для функций любого периода

Пусть $f(x)$ — периодическая функция периода $T = 2l$ (l — полупериод), кусочно-гладкая или кусочно-монотонная на отрезке $[-l, l]$. Полагая $x = at$, получим функцию $f(at)$ аргумента t , период которой равен $\frac{2\pi}{a}$. Подберем a так, чтобы период функции $f(at)$ был равен 2π , т. е.

$$\frac{T}{a} = \frac{2l}{a} = 2\pi, \text{ откуда } a = \frac{l}{\pi}.$$

Тогда подстановка $x = \frac{l}{\pi} t$ (сжатие или растяжение по оси OX) приводит к функции $f\left(\frac{lt}{\pi}\right)$ периода 2π . Эта функция удовлетворяет условию разложимости в ряд Фурье, так как она кусочно-гладкая или кусочно-монотонная на отрезке $[-\pi, \pi]$. (Нетрудно видеть, что отрезку $[-l, l]$ значений x соответствует отрезок $[-\pi, \pi]$ значений t).

Разлагая функцию $f\left(\frac{lt}{\pi}\right)$ в ряд Фурье, получим:

$$f\left(\frac{lt}{\pi}\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt), \quad (1)$$

при этом

$$\left. \begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{lt}{\pi}\right) \cos nt \, dt, \quad n=0, 1, 2, \dots, \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{lt}{\pi}\right) \sin nt \, dt, \quad n=1, 2, \dots \end{aligned} \right\} \quad (2)$$

Исходя из равенства $x = \frac{l}{\pi} t$, перейдем теперь снова к переменной x . Так как $t = \frac{\pi}{l} x$, $dt = \frac{\pi}{l} dx$ и пределам интегрирования по t от $-\pi$ до π соответствуют пределы по x от $-l$ до l , то ряд Фурье функции $f(x)$ периода $T=2l$ запишется в виде:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi}{l} x + b_n \sin \frac{n\pi}{l} x \right), \quad (3)$$

при этом

$$\begin{aligned} a_n &= \frac{1}{\pi} \cdot \frac{\pi}{l} \int_{-l}^l f(x) \cos \frac{n\pi}{l} x \, dx = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{n\pi}{l} x \, dx, \quad n=0, 1, \dots, \\ b_n &= \frac{1}{\pi} \cdot \frac{\pi}{l} \int_{-l}^l f(x) \sin \frac{n\pi}{l} x \, dx = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{n\pi}{l} x \, dx, \quad n=1, 2, \dots \end{aligned}$$

Рядом Фурье любой функции $f(x)$ периода $T=2l$, непрерывной или имеющей конечное число точек разрыва первого рода на отрезке $[-l, l]$, называется тригонометрический ряд вида:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi}{l} x + b_n \sin \frac{n\pi}{l} x \right), \quad (3')$$

коэффициенты которого определяются равенствами:

$$\left. \begin{aligned} a_0 &= \frac{1}{l} \int_{-l}^l f(x) \, dx, \\ a_n &= \frac{1}{l} \int_{-l}^l f(x) \cos \frac{n\pi}{l} x \, dx, \quad n=1, 2, \dots, \\ b_n &= \frac{1}{l} \int_{-l}^l f(x) \sin \frac{n\pi}{l} x \, dx, \quad n=1, 2, \dots \end{aligned} \right\} \quad (4)$$

Если $f(x)$ — функция периода $T=2l$, кусочно-гладкая или кусочно-монотонная на отрезке $[-l, l]$, то она разлагается в ряд Фурье, т. е. равенство (3) справедливо во всех точках непрерывности функции $f(x)$.

В точках разрыва функции $f(x)$ сумма ряда (3') равна

$$\frac{f(x-0) + f(x+0)}{2}.$$

Ряд Фурье четной и нечетной функций любого периода ($T=2l$). Пусть $f(x)$ — четная, кусочно-монотонная или кусочно-гладкая функция

на отрезке $[-l, l]$, тогда ее можно разложить на этом отрезке в ряд Фурье. По свойству интеграла по симметричному относительно $x=0$ интервалу получим $b_n=0$, $n=1, 2, \dots$,

$$\left. \begin{aligned} a_0 &= \frac{2}{l} \int_0^l f(x) dx, \\ a_n &= \frac{2}{l} \int_0^l f(x) \cos \frac{n\pi}{l} x dx, \quad n=1, 2, \dots \end{aligned} \right\} \quad (5)$$

Таким образом, ряд Фурье четной функции $f(x)$ периода $T=2l$ запишется в виде:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi}{l} x, \quad (6)$$

при этом коэффициенты ряда определяются равенствами (5).

Если $f(x)$ — нечетная функция, кусочно-монотонная или кусочно-гладкая на отрезке $[-l, l]$, то $a_n=0$, $n=0, 1, \dots$,

$$b_n = \frac{2}{l} \int_0^l f(x) \sin \frac{n\pi}{l} x dx, \quad n=1, 2, \dots \quad (7)$$

Ряд Фурье нечетной функции периода $T=2l$ имеет вид:

$$f(x) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi}{l} x, \quad (8)$$

при этом коэффициенты ряда определяются равенствами (7). Следует иметь в виду, что все сделанные раньше замечания относительно раз-

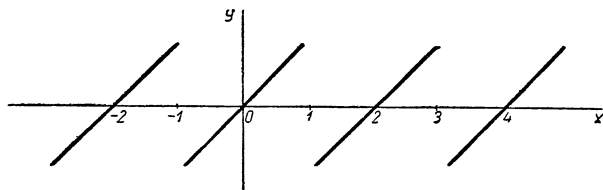


Рис. 444.

ложения в ряд Фурье функции $f(x)$, заданной на отрезке $[-\pi, \pi]$ [интервале $(-\pi, \pi)$], остаются справедливыми и для функции $f(x)$, заданной на отрезке $[-l, l]$ [интервале $(-l, l)$]. Аналогично все сказанное о разложении функции на отрезке $[0, \pi]$ [интервале $(0, \pi)$] переносится на отрезок $[0, l]$ [интервал $(0, l)$].

Пример. Разложить функцию $f(x)=x$, $-1 < x < 1$, периода $T=2$ в ряд Фурье (рис. 444).

Решение. Функция $f(x)$ — нечетная, удовлетворяющая условиям разложимости в ряд Фурье, поэтому на основании равенств (7) и (8) имеем:

$$f(x) = \sum_{n=1}^{\infty} b_n \sin n\pi x \quad (\text{здесь } l=1),$$

где

$$\begin{aligned} b_n &= 2 \int_0^1 x \sin n\pi x dx = 2 \left[-x \frac{\cos n\pi x}{n\pi} \Big|_0^1 + \frac{1}{n\pi} \int_0^1 \cos n\pi x dx \right] = \\ &= 2 \left[-\frac{\cos n\pi}{n\pi} + \frac{1}{n^2\pi^2} \sin n\pi x \Big|_0^1 \right] = -\frac{2 \cos n\pi}{n\pi} = \\ &= (-1)^{n+1} \frac{2}{n\pi}, \quad n=1, 2, 3, \dots, \end{aligned}$$

и, следовательно,

$$f(x) = \frac{2}{\pi} \left[\frac{\sin \pi x}{1} - \frac{\sin 2\pi x}{2} + \frac{\sin 3\pi x}{3} - \dots + (-1)^{n+1} \frac{\sin n\pi x}{n} + \dots \right]$$

или иначе

$$x = \frac{2}{\pi} \left[\frac{\sin \pi x}{1} - \frac{\sin 2\pi x}{2} + \frac{\sin 3\pi x}{3} - \dots + (-1)^{n+1} \frac{\sin n\pi x}{n} + \dots \right], \quad -1 < x < 1.$$

В частности, при $x = \frac{1}{2}$ получим:

$$\frac{1}{2} = \frac{2}{\pi} \left[1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \right],$$

откуда

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$$

§ 18. Ряд Фурье по любой ортогональной системе функций. Минимальное свойство коэффициентов Фурье

Функции $\varphi(x)$ и $\psi(x)$, определенные на отрезке $[a, b]$, называются *ортогональными на этом отрезке*, если

$$\int_a^b \varphi(x) \psi(x) dx = 0.$$

При этом предполагается, что интеграл в левой части равенства существует. Для простоты будем рассматривать функции, непрерывные на отрезке $[a, b]$. Последовательность функций

$$\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x), \dots, \quad (1)$$

непрерывных на отрезке $[a, b]$, называется *ортогональной системой функций на отрезке $[a, b]$* , если все функции последовательности попарно ортогональны на этом отрезке, т. е. если

$$\int_a^b \varphi_i(x) \varphi_j(x) dx = 0 \quad (i \neq j).^1)$$

Всюду дальше будем предполагать, что среди функций последовательности (1) нет функции, тождественно равной нулю.

Пример 1. Тригонометрическая система функций $\frac{1}{2}, \sin x, \cos x, \sin 2x, \cos 2x, \dots, \cos nx, \sin nx, \dots$ ортогональна на отрезке $[-\pi, \pi]$, а следовательно, в силу периодичности этих функций — и на любом отрезке длины 2π . (Это утверждение получается сразу же, если обратиться к интегралам п. 1, § 1.)

Пример 2. Система функций $\sin x, \sin 2x, \dots, \sin nx, \dots$ ортогональна на отрезке $[0, \pi]$, так как при любых натуральных m и n , $m \neq n$,

$$\int_0^\pi \sin mx \sin nx dx = \frac{1}{2} \left[\frac{\sin(m-n)x}{m-n} - \frac{\sin(m+n)x}{m+n} \right] \Big|_0^\pi = 0.$$

¹⁾ Ортогональность функций не следует понимать как перпендикулярность графиков этих функций. Например, $\sin x$ и $\cos x$ — ортогональные функции на отрезке

$[0, \pi]$, так как $\int_0^\pi \sin x \cos x dx = 0$. Обобщением понятия ортогональности функций

$\varphi(x)$ и $\psi(x)$ на отрезке $[a, b]$ является понятие *ортогональности с весом $p(x)$* этих функций на отрезке $[a, b]$, которая определяется условием

$$\int_a^b p(x) \varphi(x) \psi(x) dx = 0.$$

Заметим, что в приложениях используются и другие ортогональные системы, отличные от тригонометрических.

Ортогональность системы (1) на отрезке $[a, b]$ не нарушится, если каждую функцию системы умножить на постоянный множитель, т. е. система

$$\lambda_1 \varphi_1(x), \lambda_2 \varphi_2(x), \dots, \lambda_n \varphi_n(x), \dots \quad (1')$$

также ортогональна на отрезке $[a, b]$. Действительно,

$$\int_a^b \lambda_i \varphi_i(x) \lambda_j \varphi_j(x) dx = \lambda_i \lambda_j \int_a^b \varphi_i(x) \varphi_j(x) dx = 0 \quad (i \neq j).$$

Система (1) называется *ортогональной и нормированной (ортонормированной)* на отрезке $[a, b]$, если она ортогональна на этом отрезке, т. е.

$$\int_a^b \varphi_i(x) \varphi_j(x) dx = 0, \quad i \neq j, \quad i, j = 1, 2, \dots,$$

и нормирована, т. е.

$$\int_a^b [\varphi_i(x)]^2 dx = 1, \quad i = 1, 2, \dots,$$

или, что то же, если

$$\int_a^b \varphi_i(x) \varphi_j(x) dx = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \quad (2)$$

Например, система функций

$$\sqrt{\frac{2}{\pi}} \sin x, \sqrt{\frac{2}{\pi}} \sin 2x, \dots, \sqrt{\frac{2}{\pi}} \sin nx, \dots$$

ортонормированная на отрезке $[0, \pi]$.

В самом деле, эта система ортогональна (пример 2) и нормирована, так как

$$\frac{2}{\pi} \int_0^{\pi} \sin^2 nx dx = 1, \quad n = 1, 2, \dots$$

Заметим, что любую ортогональную на отрезке $[a, b]$ систему (1) можно нормировать, если умножить каждую функцию этой системы на некоторый множитель, называемый *нормирующим*.

В самом деле, система функций

$$\lambda_1 \varphi_1(x), \lambda_2 \varphi_2(x), \dots, \lambda_n \varphi_n(x), \dots \quad (1')$$

ортогональна на отрезке $[a, b]$. Подберем теперь множители λ_n ($n = 1, 2, \dots$) так, чтобы выполнялось условие

$$\int_a^b [\lambda_n \varphi_n(x)]^2 dx = 1, \quad n = 1, 2, \dots$$

Из этого условия определим

$$\lambda_n = \frac{1}{\pm \sqrt{\int_a^b [\varphi_n(x)]^2 dx}}, \quad n = 1, 2, \dots \quad (3)$$

Эти числа λ_n ($n = 1, 2, \dots$) называются *нормирующими множителями* (знак у корня можно брать любой), при таких λ_n ($n = 1, 2, \dots$) система (1') будет ортонормированной.

Рядом по ортогональной системе (1) называется ряд вида:

$$\sum_{n=1}^{\infty} a_n \varphi_n(x),$$

где числа a_n ($n=1, 2, \dots$) — коэффициенты ряда. Пусть функция $f(x)$ является суммой равномерно сходящегося на отрезке $[a, b]$ ряда по ортогональной системе (1):

$$f(x) = a_1 \varphi_1(x) + a_2 \varphi_2(x) + \dots + a_n \varphi_n(x) + \dots$$

Определим коэффициенты этого ряда. Функция $f(x)$, как сумма равномерно сходящегося на отрезке $[a, b]$ ряда непрерывных функций, является непрерывной функцией на этом отрезке. Умножая ряд на функцию $\varphi_n(x)$ (при этом равномерная сходимость ряда не нарушится) и интегрируя его почленно в пределах от a до b , получим:

$$\begin{aligned} \int_a^b f(x) \varphi_n(x) dx &= a_1 \int_a^b \varphi_1(x) \varphi_n(x) dx + \\ &+ a_2 \int_a^b \varphi_2(x) \varphi_n(x) dx + \dots + a_n \int_a^b [\varphi_n(x)]^2 dx + \\ &+ \dots = a_n \int_a^b [\varphi_n(x)]^2 dx, \end{aligned}$$

откуда

$$a_n = \frac{\int_a^b f(x) \varphi_n(x) dx}{\int_a^b [\varphi_n(x)]^2 dx}, \quad n=1, 2, \dots \quad (4)$$

Если система (1) ортонормированная, то в этом случае

$$a_n = \int_a^b f(x) \varphi_n(x) dx, \quad n=1, 2, \dots \quad (4')$$

Пусть теперь $f(x)$ — любая функция, непрерывная или имеющая конечное число точек разрыва первого рода на отрезке $[a, b]$. Рядом Фурье такой функции $f(x)$ на отрезке $[a, b]$ по ортогональной системе (1) называется ряд:

$$\sum_{n=1}^{\infty} a_n \varphi_n(x), \quad (5)$$

коэффициенты которого определяются равенствами (4):

$$f(x) \sim \sum_{n=1}^{\infty} a_n \varphi_n(x)^1.$$

Если ряд Фурье функции $f(x)$ по системе (1) сходится к функции $f(x)$ в каждой ее точке непрерывности, принадлежащей отрезку $[a, b]$, то знак соответствия заменяется знаком равенства, $f(x) = \sum_{n=1}^{\infty} a_n \varphi_n(x)$ на отрезке $[a, b]$. В этом случае говорят, что функция $f(x)$ на отрезке $[a, b]$ разлагается по ортогональной системе (1).

¹⁾ Заметим, что ряд Фурье функции $f(x)$ по данной ортогональной системе и по соответствующей ей ортонормированной системе один и тот же.

Заметим, что если рассмотреть ортогональную на отрезке $[-\pi, \pi]$ тригонометрическую систему функций:

$$\frac{1}{2}, \cos x, \sin x, \dots, \cos nx, \sin nx, \dots$$

и любую функцию $f(x)$, непрерывную или имеющую конечное число точек разрыва первого рода на отрезке $[-\pi, \pi]$, то рядом Фурье функции $f(x)$ по тригонометрической системе является тригонометрический ряд

$$\frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \dots,$$

коэффициенты которого на основании формул (4) определяются равенствами:

$$a_0 = \frac{\int_{-\pi}^{\pi} f(x) \cdot \frac{1}{2} dx}{\int_{-\pi}^{\pi} \left(\frac{1}{2}\right)^2 dx} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx,$$

$$a_n = \frac{\int_{-\pi}^{\pi} f(x) \cos nx dx}{\int_{-\pi}^{\pi} \cos^2 nx dx} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx,$$

$$b_n = \frac{\int_{-\pi}^{\pi} f(x) \sin nx dx}{\int_{-\pi}^{\pi} \sin^2 nx dx} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx.$$

Понятие о среднем квадратическом уклонении функций. Пусть $f(x)$,

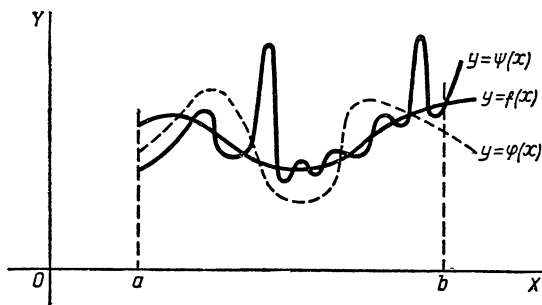


Рис. 445.

$\varphi(x)$, $\psi(x)$ — непрерывные функции на отрезке $[a, b]$, их графики изображены на рис. 445.

Прежде чем ответить на вопрос, какая из двух функций: $\varphi(x)$ или $\psi(x)$ — лучше приближает функцию $f(x)$ на отрезке $[a, b]$, нужно сначала сказать, что принимается за характеристику степени приближения функций на рассматриваемом отрезке.

Если близость функции $f_1(x)$ к функции $f(x)$ на отрезке $[a, b]$ характеризовать наибольшей величиной модуля разности $|f(x) - f_1(x)|$ на этом отрезке, то в этом смысле функция $\varphi(x)$ дает лучшее, по сравнению с функцией $\psi(x)$, приближение функции $f(x)$. С приближением такого вида мы встречались при рассмотрении равномерной сходимости последовательности функций на отрезке $[a, b]$ (там рассматривался на отрезке $[a, b]$ модуль разности функций $|f(x) - f_n(x)|$, $f(x) = \lim_{n \rightarrow \infty} f_n(x)$).

Если рассматривать, образно выражаясь, «наиболее вероятное» приближение функции $f_1(x)$ к функции $f(x)$ на отрезке $[a, b]$, отбрасывая некоторые узкие промежутки значений x , на которых отклонение в определен-

ном выше смысле одной функции от другой может быть и значительным, то тогда функция $\psi(x)$ дает лучшее, по сравнению с функцией $\varphi(x)$, приближение функции $f(x)$ на этом отрезке.

Определим теперь, что будем понимать под близостью функций в этом новом смысле.

Пусть на отрезке $[a, b]$ имеем две непрерывные функции $f(x)$ и $\varphi(x)$. Разделим этот отрезок на n равных частей. Точками x_i ($i = 0, 1, \dots, n$), при этом $x_0 = a$, $x_n = b$. Разность $f(x_i) - \varphi(x_i)$ является ошибкой при замене функции $f(x)$ функцией $\varphi(x)$ в точке x_i . Число δ_n , определяемое равенством

$$\delta_n^2 = \frac{1}{n} \sum_{i=1}^n [f(x_i) - \varphi(x_i)]^2,$$

называется *средней квадратической ошибкой* (или средним квадратическим отклонением) при замене функции $f(x)$ функцией $\varphi(x)$ или, наоборот, в рассматриваемых точках x_i , $i = 1, 2, \dots, n$ ¹⁾. Предел среднего квадратического отклонения при неограниченном измельчении отрезка $[a, b]$ называется *средним квадратическим отклонением функции* $\varphi(x)$ от функции $f(x)$ (или наоборот) на отрезке $[a, b]$. Среднее квадратическое отклонение будем обозначать через δ , тогда по определению

$$\delta^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n [f(x_i) - \varphi(x_i)]^2.$$

Пусть $\frac{b-a}{n} = \Delta x_i$ ($i = 1, 2, \dots, n$), тогда $\frac{1}{n} = \frac{\Delta x_i}{b-a}$ и δ^2 запишется

в виде:

$$\delta^2 = \frac{1}{b-a} \lim_{n \rightarrow \infty} \sum_{i=1}^n [f(x_i) - \varphi(x_i)]^2 \Delta x_i.$$

Так как предел суммы, записанной в правой части, равен интегралу $\int_a^b [f(x) - \varphi(x)]^2 dx$, то

$$\delta^2 = \frac{1}{b-a} \int_a^b [f(x) - \varphi(x)]^2 dx.$$

Заметим, что если $|f(x) - \varphi(x)| \leq m$, для всех x отрезка $[a, b]$, то и среднее квадратическое отклонение $\delta \leq m$.

На основании выражения δ^2 через интеграл можно заключить, что среднее квадратическое отклонение функции $\varphi(x)$ от функции $f(x)$ может быть мало и в том случае, когда обычное отклонение функции $\varphi(x)$ от функции $f(x)$ значительное, но на небольших промежутках.

Минимальное свойство коэффициентов Фурье. Пусть

$$\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x), \dots \quad (6)$$

¹⁾ Так как ошибки могут быть как положительные, так и отрицательные, то удобнее рассматривать квадраты ошибок, которые всегда неотрицательны и, следовательно, при суммировании не могут взаимно уничтожаться.

— ортогональная система функций на отрезке $[a, b]$. Выражение $c_1 \varphi_1(x) + c_2 \varphi_2(x) + \dots + c_n \varphi_n(x) = \sum_{k=1}^n c_k \varphi_k(x)$, где c_k ($k=1, 2, \dots, n$) — числа, называется *полиномом n -го порядка относительно заданной системы функций*.

Пусть функция $f(x)$ на отрезке $[a, b]$ непрерывна или имеет конечное число точек разрыва 1-го рода. Рассмотрим следующую задачу.

Среди всех полиномов $P_n(x) = \sum_{k=1}^n c_k \varphi_k(x)$ порядка n найти тот, который дает наименьшее среднее квадратическое отклонение от функции $f(x)$ на отрезке $[a, b]$ или, другими словами, найти полином, который дает наилучшее среднее квадратическое приближение функции $f(x)$ на отрезке $[a, b]$.

Не нарушая общности рассуждений можно считать, что система функций $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x), \dots$ — ортонормированная на рассматриваемом отрезке, последнее всегда можно сделать, если умножить функции системы на нормирующие множители. Итак, задача состоит в определении коэффициентов c_k ($k=1, 2, \dots, n$) полинома $P_n(x)$, при которых интеграл

$$\int_a^b [f(x) - P_n(x)]^2 dx = \int_a^b [f(x) - \sum_{k=1}^n c_k \varphi_k(x)]^2 dx$$

имеет наименьшее значение¹⁾.

Преобразуем этот интеграл:

$$\begin{aligned} \int_a^b [f(x) - \sum_{k=1}^n c_k \varphi_k(x)]^2 dx &= \int_a^b \{ [f(x)]^2 - 2f(x) \sum_{k=1}^n c_k \varphi_k(x) + [\sum_{k=1}^n c_k \varphi_k(x)]^2 \} dx = \\ &= \int_a^b [f(x)]^2 dx - 2 \sum_{k=1}^n c_k \int_a^b f(x) \varphi_k(x) dx + \\ &+ \sum_{k=1}^n c_k^2 \int_a^b [\varphi_k(x)]^2 dx + 2 \sum_{i < k} c_i c_k \int_a^b \varphi_i(x) \varphi_k(x) dx. \end{aligned}$$

Учитывая, что при любом k интеграл $\int_a^b f(x) \varphi_k(x) dx = a_k$, a_k — коэффициент Фурье функции $f(x)$ по данной ортонормированной системе, а также что

$$\int_a^b \varphi_i(x) \varphi_k(x) dx = 0, \quad i \neq k, \quad i, k = 1, 2, \dots, n,$$

и

$$\int_a^b [\varphi_k(x)]^2 dx = 1, \quad k = 1, 2, \dots, n$$

(по определению ортонормированной системы), получим:

$$\begin{aligned} \int_a^b [f(x) - \sum_{k=1}^n c_k \varphi_k(x)]^2 dx &= \int_a^b [f(x)]^2 dx - 2 \sum_{k=1}^n c_k a_k + \\ &+ \sum_{k=1}^n c_k^2 = \int_a^b [f(x)]^2 dx + \sum_{k=1}^n (c_k - a_k)^2 - \sum_{k=1}^n a_k^2. \end{aligned}$$

¹⁾ Рассматриваемый интеграл только на постоянный множитель отличается от δ^2 .

²⁾ Запись $\sum_{i < k} c_i c_k \int_a^b \varphi_i(x) \varphi_k(x) dx$ означает, что суммирование ведется по всем индексам i, k , удовлетворяющим условию $i < k$.

Нетрудно видеть, что интеграл $\int_a^b [f(x) - \sum_{k=1}^n c_k \varphi_k(x)]^2 dx$ будет минимальным, когда $\sum_{k=1}^n (c_k - a_k)^2 = 0$, т. е. при $c_k = a_k$, $k = 1, 2, \dots, n$.

Таким образом, из всех полиномов n -го порядка относительно данной ортогональной системы на отрезке $[a, b]$ наименьшее среднее квадратическое отклонение от заданной функции $f(x)$ на этом отрезке имеет полином с коэффициентами Фурье функции $f(x)$ относительно данной ортогональной системы, т. е. n -я частная сумма ряда Фурье функции $f(x)$ по этой ортогональной системе.

Таким образом, независимо от того, сходится ряд Фурье функции $f(x)$ по ортогональной системе на отрезке $[a, b]$ к этой функции или нет, его n -я частная сумма дает наилучшее среднее квадратическое приближение функции $f(x)$ на этом отрезке по сравнению со всеми другими полиномами порядка n относительно данной ортогональной системы.

Для ортогональной на отрезке $[-\pi, \pi]$ тригонометрической системы:

$$\frac{1}{2}, \cos x, \sin x, \dots, \cos nx, \sin nx, \dots$$

тригонометрическим полиномом n -го порядка называется функция вида:

$$\frac{c_0}{2} + \sum_{k=1}^n (c_k \cos kx + d_k \sin kx).$$

Из всех тригонометрических полиномов порядка n наименьшее среднее квадратическое отклонение на отрезке $[-\pi, \pi]$ от заданной на этом отрезке функции $f(x)$ имеет полином с коэффициентами Фурье этой функции, т. е. полином

$$\frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx).$$

Из проведенных выше рассуждений получаем, что

$$\begin{aligned} \min \int_a^b [f(x) - \sum_{k=1}^n c_k \varphi_k(x)]^2 dx &= \int_a^b [f(x) - \\ &- \sum_{k=1}^n a_k \varphi_k(x)]^2 dx = \int_a^b [f(x)]^2 dx - \sum_{k=1}^n a_k^2, \end{aligned}$$

откуда, если учесть, что подинтегральная функция неотрицательна, имеем:

$$\int_a^b [f(x)]^2 dx - \sum_{k=1}^n a_k^2 \geq 0$$

и, значит,

$$\sum_{k=1}^n a_k^2 \leq \int_a^b [f(x)]^2 dx.$$

Из этого неравенства, справедливого при любом n , следует сходимость ряда $\sum_{k=1}^{\infty} a_k^2$. Переходя в этом неравенстве к пределу при $n \rightarrow \infty$, получаем:

$$\sum_{k=1}^{\infty} a_k^2 \leq \int_a^b [f(x)]^2 dx$$

неравенство Бесселя¹⁾.

¹⁾ Бессель (1784 — 1846) — немецкий математик и астроном.

Нетрудно видеть, что неравенство Бесселя в случае ортогональной (но не ортонормированной) системы принимает вид:

$$\sum_{k=1}^{\infty} \frac{a_k^2}{\lambda_k^2} \leq \int_a^b [f(x)]^2 dx,$$

где λ_k — нормирующий множитель для функции $\varphi_k(x)$.

Если ряд Фурье функции $f(x)$ по ортонормированной системе на отрезке $[a, b]$ равномерно сходится к функции $f(x)$ на этом отрезке, то

$$\int_a^b [f(x) - \sum_{k=1}^n a_k \varphi_k(x)]^2 dx \rightarrow 0$$

и неравенство Бесселя превращается в равенство:

$$\sum_{k=1}^{\infty} a_k^2 = \int_a^b [f(x)]^2 dx.$$

В случае любой ортогональной системы это равенство принимает вид:

$$\sum_{k=1}^{\infty} \frac{a_k^2}{\lambda_k^2} = \int_a^b [f(x)]^2 dx.$$

Заметим, что из сходимости ряда $\sum_{k=1}^{\infty} a_k^2$ непосредственно следует, что $a_k \rightarrow 0$, т. е. коэффициенты Фурье функции $f(x)$ по любой ортонормированной системе стремятся к нулю.

§ 19. Комплексная форма ряда Фурье

Пусть $f(x)$ — периодическая функция периода $T = 2\pi$, удовлетворяющая условиям разложимости в ряд Фурье, тогда

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

при этом коэффициенты ряда определяются равенствами (2) § 16. Преобразуем общий член ряда $a_n \cos nx + b_n \sin nx$ с помощью формул Эйлера:

$$\begin{aligned} a_n \cos nx + b_n \sin nx &= a_n \frac{e^{inx} + e^{-inx}}{2} + b_n \frac{e^{inx} - e^{-inx}}{2i} = a_n \frac{e^{inx} + e^{-inx}}{2} - \\ &- ib_n \frac{e^{inx} - e^{-inx}}{2} = \frac{a_n - ib_n}{2} e^{inx} + \frac{a_n + ib_n}{2} e^{-inx}. \end{aligned}$$

Если положить:

$$\begin{aligned} c_0 &= \frac{a_0}{2}, \\ c_n &= \frac{a_n - ib_n}{2}, \\ c_{-n} &= \frac{a_n + ib_n}{2} \end{aligned}$$

или, что то же,

$$c_{-n} = \overline{c_n},$$

то общий член ряда Фурье запишется в виде:

$$a_n \cos nx + b_n \sin nx = c_n e^{inx} + c_{-n} e^{-inx}$$

и, таким образом, частная сумма ряда Фурье запишется так:

$$\frac{a_0}{2} + \sum_{n=1}^N (a_n \cos nx + b_n \sin nx) = c_0 + \sum_{n=1}^N (c_n e^{inx} + c_{-n} e^{-inx}) = \sum_{-N}^N c_n e^{inx}.$$

Переходя в этом равенстве к пределу при $N \rightarrow +\infty$ и обозначив

$$\lim_{N \rightarrow +\infty} \sum_{-N}^N c_n e^{inx} = \sum_{-\infty}^{+\infty} c_n e^{inx}.$$

получим:

$$f(x) = \lim_{N \rightarrow \infty} \left[\frac{a_0}{2} + \sum_{n=1}^N (a_n \cos nx + b_n \sin nx) \right] = \lim_{N \rightarrow \infty} \sum_{-N}^N c_n e^{inx} = \sum_{-\infty}^{+\infty} c_n e^{inx},$$

т. е.

$$f(x) = \sum_{-\infty}^{+\infty} c_n e^{inx}. \quad (1)$$

Найдем теперь выражения для коэффициентов c_n . Действительно, если учесть выражения для a_n и b_n , то получим:

$$\begin{aligned} c_n &= \frac{a_n - ib_n}{2} = \frac{1}{2} \left(\frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx - i \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \right) = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) (\cos nx - i \sin nx) \, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx, \end{aligned}$$

т. е.

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx \quad (n = 1, 2, \dots),$$

$$c_0 = \frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx,$$

$$\begin{aligned} c_{-n} &= \frac{a_n + ib_n}{2} = \frac{1}{2} \left(\frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx + i \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \right) = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) (\cos nx + i \sin nx) \, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{inx} \, dx \quad (n = 1, 2, \dots). \end{aligned}$$

Нетрудно видеть, что при всех целых n справедливо равенство

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx \quad (n = 0, \pm 1, \pm 2, \dots), \quad (2)$$

при этом интегрирование можно вести по любому отрезку длины 2π .

Выражение $\sum_{-\infty}^{+\infty} c_n e^{inx}$ называется *комплексной формой ряда Фурье функции $f(x)$* .

Пример. 1) Записать ряд Фурье в комплексной форме для периодической функции $f(x)$ периода $T = 2\pi$, определенной при $0 < x < 2\pi$ равенством $f(x) = e^x$.

¹⁾ Совершенно так же, как для действительных функций действительной переменной, для комплексных функций действительной переменной вводятся понятия производной и интегралов, неопределенного и определенного. Можно показать, что формулы дифференцирования и интегрирования, а также основные свойства определенного интеграла остаются справедливыми и в этом случае.

2) От комплексной формы перейти к действительной форме ряда Фурье функции $f(x)$.

Решение. 1) По формулам (1), (2) запишем: $f(x) = \sum_{-\infty}^{+\infty} c_n e^{inx}$, при этом

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx = \frac{1}{2\pi} \int_0^{2\pi} e^x \cdot e^{-inx} dx = \frac{1}{2\pi} \int_0^{2\pi} e^{(1-in)x} dx = \\ &= \frac{1}{2\pi} \frac{1}{1-in} e^{(1-in)x} \Big|_0^{2\pi} = \frac{1}{2\pi} \cdot \frac{1}{1-in} (e^{2\pi} - 1) \quad (n=0, \pm 1, \pm 2, \dots) \end{aligned}$$

или, иначе,

$$c_n = \frac{e^{2\pi} - 1}{2\pi} \frac{1 + in}{1 + n^2} \quad (n=0, \pm 1, \pm 2, \dots).$$

Подставляя c_n ($n=0, \pm 1, \dots$) в ряд для $f(x)$, получим ряд Фурье этой функции

$$f(x) = \frac{e^{2\pi} - 1}{2\pi} \sum_{-\infty}^{+\infty} \frac{1 + in}{1 + n^2} e^{inx}.$$

Отсюда следует, что

$$e^x = \frac{e^{2\pi} - 1}{2\pi} \sum_{-\infty}^{+\infty} \frac{1 + in}{1 + n^2} e^{inx}, \quad 0 < x < 2\pi.$$

2) Если учесть, что

$$\begin{aligned} a_n &= 2\operatorname{Re}(c_n), & n=0, 1, 2, \dots, \\ b_n &= -2\operatorname{Im}(c_n), & n=1, 2, \dots, \end{aligned}$$

то найдем:

$$\begin{aligned} a_n &= \frac{e^{2\pi} - 1}{\pi} \frac{1}{1 + n^2}, & n=0, 1, 2, \dots, \\ b_n &= -\frac{e^{2\pi} - 1}{\pi} \frac{n}{1 + n^2}, & n=1, 2, \dots \end{aligned}$$

и, значит,

$$f(x) = \frac{e^{2\pi} - 1}{\pi} \left[\frac{1}{2} + \sum_{n=1}^{\infty} \left(\frac{\cos nx}{1 + n^2} - \frac{n \sin nx}{1 + n^2} \right) \right].$$

Или, в другом виде,

$$e^x = \frac{e^{2\pi} - 1}{\pi} \left[\frac{1}{2} + \sum_{n=1}^{\infty} \left(\frac{\cos nx}{1 + n^2} - \frac{n \sin nx}{1 + n^2} \right) \right] \quad \text{при } 0 < x < 2\pi.$$

Комплексная форма ряда Фурье для периодической функции периода $T=2l$. Пусть $f(x)$ — периодическая функция периода $T=2l$, удовлетворяет условиям разложимости в ряд Фурье. Тогда подстановка $x = \frac{l}{\pi} t$ приводит к функции $f\left(\frac{lt}{\pi}\right)$ с периодом 2π , разложимой в ряд Фурье. Для такой функции по формулам (1) и (2) имеем:

$$f\left(\frac{lt}{\pi}\right) = \sum_{-\infty}^{+\infty} c_n e^{int}, \quad \text{где } c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(\frac{lt}{\pi}\right) e^{-int} dt.$$

Переходя к аргументу x , с помощью подстановки $t = \frac{\pi x}{l}$ получим:

$$f(x) = \sum_{-\infty}^{+\infty} c_n e^{i \frac{n\pi}{l} x}, \quad (3)$$

при этом

$$c_n = \frac{1}{2\pi} \cdot \frac{\pi}{l} \int_{-l}^l f(x) e^{-i \frac{n\pi}{l} x} dx$$

или окончательно

$$c_n = \frac{1}{2l} \int_{-l}^l f(x) e^{-i \frac{n\pi}{l} x} dx. \quad (4)$$

Выражение $\sum_{-\infty}^{+\infty} c_n e^{i \frac{n\pi}{l} x}$ называется *комплексной формой ряда Фурье* этой функции $f(x)$.

§ 20. Задача о колебании струны

Рассмотрим применение рядов Фурье к решению задачи о колебании струны.

Пусть в состоянии равновесия натянутая струна длины l закреплена на концах.

Примем за ось OX направление струны и пусть $x=0$ и $x=l$ — ее концы.

Предположим, что струна выведена из состояния равновесия и каждой ее точке сообщена некоторая начальная скорость, а затем ей предоставлена возможность колебаться свободно. Будем рассматривать малые колебания струны (длина струны, а следовательно, и ее натяжение постоянны), происходящие в вертикальной плоскости (т. е. рассматривать смещение каждой точки струны только в вертикальном направлении).

При сделанных выше допущениях можно показать, что функция $u(x, t)$, характеризующая положение струны в каждый момент времени t , удовлетворяет уравнению¹⁾

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad (1)$$

Рис. 446.

где a — положительное число. (На рис. 446 изображен график функции $u(x, t)$ при некотором значении t , он дает форму струны в момент времени t .)

Уравнение (1) является линейным однородным уравнением с частными производными второго порядка. Наша задача — найти функцию $u(x, t)$, график которой дает форму струны в любой момент времени t , т. е. найти решение уравнения (1) при граничных условиях:

$$u(0, t) = 0, \quad u(l, t) = 0, \quad (2)$$

выражающих тот факт, что струна закреплена на концах, и начальных условиях:

$$u(x, 0) = \varphi(x), \quad u'_t(x, 0) = \psi(x) \quad (3)$$

[функции $\varphi(x)$ и $\psi(x)$ заданы]. Первое из условий (3) характеризует начальное положение струны, второе — характеризует начальные скорости ее точек.

Сначала будем искать решения уравнения (1), удовлетворяющие граничным условиям (2). Нетрудно видеть, что $u(x, t) \equiv 0$ является ре-

¹⁾ Уравнение (1) мы приводим без вывода. Оно выводится в различных курсах механики, в более полных курсах математического анализа и в специальных курсах по уравнениям математической физики (см. Н. С. Кошляков и др. Основные дифференциальные уравнения математической физики).

шением уравнения (1), удовлетворяющим условиям (2). Будем искать решения, не равные тождественно нулю, представимые в виде произведения

$$u(x, t) = X(x) T(t), \quad (4)$$

где $X(x)$ и $T(t)$ — дважды непрерывно дифференцируемые функции, соответственно зависящие от x , $0 \leq x \leq l$, и t , $-\infty < t < +\infty$. Так как функция $u(x, t)$ не обращается тождественно в нуль, то и функции $X(x)$, $T(t)$ также не обращаются тождественно в нуль.

Подстановка выражения (4) в уравнение (1) дает:

$$XT'' = a^2 X''T,$$

откуда

$$\frac{X''}{X} = \frac{T''}{a^2 T}.$$

Левая часть этого равенства не зависит от t , правая — не зависит от x , следовательно, величина этих отношений — постоянная, обозначим ее через λ :

$$\frac{X''}{X} = \frac{T''}{a^2 T} = \lambda.$$

Из полученного соотношения имеем:

$$X'' - \lambda X = 0, \quad (5)$$

$$T'' - \lambda a^2 T = 0. \quad (6)$$

Уравнения (5), (6) — обыкновенные линейные однородные дифференциальные уравнения второго порядка с постоянными коэффициентами, которым удовлетворяют искомые функции $X(x)$ и $T(t)$.

Как следует из условий (2),

$$u(0, t) = X(0) T(t) = 0, \text{ откуда } X(0) = 0,$$

$$u(l, t) = X(l) T(t) = 0, \text{ откуда } X(l) = 0$$

(если учесть, что функция $T(t)$ не обращается тождественно в нуль). Таким образом, наша задача свелась к отысканию решений уравнений (5) и (6), не обращающихся тождественно в нуль. При этом функция $X(x)$ должна удовлетворять еще условиям:

$$X(0) = 0, \quad X(l) = 0. \quad (7)$$

Используя условия (7), докажем, что λ в уравнениях (5) и (6) отрицательное число.

Разберем все возможные случаи.

а) Пусть $\lambda = 0$. Тогда уравнение (5) примет вид: $X'' = 0$ и его общее решение запишется так:

$$X(x) = C_1 x + C_2.$$

Подчинив $X(x)$ условию (7), получим:

$$X(0) = C_2 = 0,$$

$$X(l) = C_1 l = 0,$$

откуда $C_1 = C_2 = 0$ и, следовательно, $X(x) \equiv 0$, что невозможно, так как мы рассматриваем решения, не обращающиеся тождественно в нуль.

б) Пусть $\lambda > 0$. Тогда, решив уравнение

$$X'' - \lambda X = 0,$$

$$r^2 - \lambda = 0, \quad r_{1,2} = \pm \sqrt{\lambda},$$

$$\text{получим } X(x) = C_1 e^{\sqrt{\lambda} x} + C_2 e^{-\sqrt{\lambda} x},$$

и, подчинив $X(x)$ условию (7), найдем, что $X(x) \equiv 0$.

Действительно, $X(0) = C_1 + C_2 = 0$, откуда $C_2 = -C_1$,

$$X(l) = C_1 e^{\sqrt{\lambda} l} - C_1 e^{-\sqrt{\lambda} l} = C_1 \frac{e^{2\sqrt{\lambda} l} - 1}{e^{\sqrt{\lambda} l}} = 0,$$

а так как $l \neq 0$, то последнее равенство возможно лишь при $C_1 = 0$, а значит, $C_2 = 0$ и $X(x) \equiv 0$, а это невозможно.

в) Остается только последняя возможность: $\lambda < 0$. Если положить $\lambda = -k$, $k > 0$, то уравнения (5) и (6) запишутся в виде:

$$\begin{aligned} X'' + kX &= 0, \\ T'' + ka^2 T &= 0. \end{aligned}$$

Решая эти линейные однородные уравнения, для которых соответствующие характеристические уравнения имеют корни:

$$r_{1,2} = \pm i \sqrt{k} \text{ и } r_{1,2} = \pm ia \sqrt{k},$$

получим:

$$\begin{aligned} X(x) &= C_1 \cos \sqrt{k} x + C_2 \sin \sqrt{k} x, \\ T(t) &= D_1 \cos a \sqrt{k} t + D_2 \sin a \sqrt{k} t, \end{aligned}$$

где C_1, C_2, D_1, D_2 — произвольные постоянные. Подчинив $X(x)$ условиям (7), найдем:

$$X(0) = C_1 = 0, \quad X(l) = C_2 \sin \sqrt{k} l = 0,$$

откуда [так как $C_2 \neq 0$, в силу того что $X(x)$ не обращается тождественно в нуль] $\sin \sqrt{k} l = 0$, т. е.

$$\sqrt{k} l = n\pi \quad (n = 1, 2, \dots)$$

и, значит,

$$k = \left(\frac{n\pi}{l} \right)^2 \quad (n = 1, 2, \dots).$$

Учитывая это, можно записать:

$$\left. \begin{aligned} X(x) &= C_2 \sin \frac{n\pi}{l} x, \\ T(t) &= D_1 \cos \frac{an\pi}{l} t + D_2 \sin \frac{an\pi}{l} t \end{aligned} \right\} \quad (n = 1, 2, \dots)$$

и, следовательно,

$$\begin{aligned} u(x, t) &= X(x) T(t) = \\ &= \left(A \cos \frac{an\pi}{l} t + B \sin \frac{an\pi}{l} t \right) \sin \frac{n\pi}{l} x, \quad (n = 1, 2, \dots), \end{aligned} \quad (8)$$

где $A = C_2 D_1$, $B = C_2 D_2$.

Наоборот, при любом целом неотрицательном n и любых A и B соответствующая функция $u(x, t)$, определяемая равенством (8), удовлетворяет уравнению (1) (что легко проверить, подставляя эту функцию в уравнение) и начальным условиям (2).

Так как A и B любые и, вообще говоря, различные для различных значений n , то будем обозначать их соответственно через A_n и B_n . При этих условиях любое решение (8) запишется в виде:

$$u_n(x, t) = \left(A_n \cos \frac{an\pi}{l} t + B_n \sin \frac{an\pi}{l} t \right) \sin \frac{n\pi}{l} x \quad (n = 1, 2, \dots), \quad (8')$$

где A_n и B_n — произвольные постоянные.

Попытаемся определить постоянные A_n и B_n ($n=1, 2, \dots$) таким образом, чтобы ряд

$$u(x, t) = \sum_{n=1}^{\infty} \left(A_n \cos \frac{an\pi}{l} t + B_n \sin \frac{an\pi}{l} t \right) \sin \frac{n\pi}{l} x \quad (9)$$

удовлетворял уравнению (1), граничным условиям (2) и начальным условиям (3).

Если ряд (9) допускает почленное двукратное дифференцирование по переменным x и t , то функция $u(x, t)$, определяемая этим рядом, является решением уравнения (1), что легко проверить непосредственной подстановкой в уравнение. Функция $u(x, t)$ удовлетворяет и условиям (2).

Поэтому наша задача свелась к определению таких коэффициентов A_n и B_n ряда (9), которые бы обеспечили возможность почленного двукратного дифференцирования этого ряда и таких, чтобы функция $u(x, t)$, определяемая равенством (9), удовлетворяла начальным условиям (3).

Итак, подчиним функцию $u(x, t)$ начальным условиям (3), т. е. подберем A_n и B_n так, чтобы выполнялись условия

$$u(x, 0) = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi}{l} x = \varphi(x), \quad (10)$$

$$u_t(x, 0) = \sum_{n=1}^{\infty} \frac{an\pi}{l} B_n \sin \frac{n\pi}{l} x = \psi(x). \quad (11)$$

Равенства (10), (11) являются соответственно разложениями функций $\varphi(x)$ и $\psi(x)$ на отрезке $[0, l]$ в ряды Фурье по синусам. (При этом предполагается, что функции $\varphi(x)$ и $\psi(x)$ удовлетворяют условиям разложимости в ряды Фурье на этом отрезке).

Тогда по формуле (7) § 17 коэффициенты рядов (10) и (11) запишутся так:

$$\left. \begin{aligned} A_n &= \frac{2}{l} \int_0^l \varphi(x) \sin \frac{n\pi}{l} x dx, \\ \frac{an\pi}{l} B_n &= \frac{2}{l} \int_0^l \psi(x) \sin \frac{n\pi}{l} x dx. \end{aligned} \right\} (n=1, 2, \dots)$$

Таким образом, решение задачи о колебании струны с заданными граничными и начальными условиями дается формулой

$$u(x, t) = \sum_{n=1}^{\infty} \left(A_n \cos \frac{an\pi}{l} t + B_n \sin \frac{an\pi}{l} t \right) \sin \frac{n\pi}{l} x, \quad (12)$$

где

$$\left. \begin{aligned} A_n &= \frac{2}{l} \int_0^l \varphi(x) \sin \frac{n\pi}{l} x dx, \\ B_n &= \frac{2}{an\pi} \int_0^l \psi(x) \sin \frac{n\pi}{l} x dx. \end{aligned} \right\} (n=1, 2, \dots)$$

Не останавливаясь на выяснении условий, которые достаточно наложить на функции $\varphi(x)$ и $\psi(x)$, чтобы при соответствующих значениях коэффициентов A_n и B_n обеспечивалась возможность двукратного почленного дифференцирования ряда (9) или, что то же, ряда (12) по переменным x и t , заметим только: если функция $\varphi(x)$ на отрезке $[0, l]$ имеет непрерывные производные до третьего порядка включительно и,

кроме того, $\varphi(0) = \varphi(l) = 0$, $\varphi''(0) = \varphi''(l) = 0$, а функция $\psi(x)$ на отрезке $[0, l]$ имеет непрерывные производные до второго порядка включительно и, кроме того, $\psi(0) = \psi(l) = 0$, то почленное двукратное дифференцирование ряда (12) по переменным x и t может быть обосновано.

§ 21. Интеграл Фурье

Пусть функция $f(x)$ задана на всей числовой оси и на каждом отрезке $[-l, l]$ (l — любое число) она кусочно-гладкая или кусочно-монотонная¹⁾. Кроме того, предположим, что $f(x)$ — абсолютно интегрируемая функция на всей числовой оси, т. е. несобственный интеграл

$$\int_{-\infty}^{+\infty} |f(x)| dx$$

сходится. Как известно, такую функцию $f(x)$ на каждом интервале $(-l, l)$ можно разложить в ряд Фурье:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi}{l} x + b_n \sin \frac{n\pi}{l} x \right), \quad |x| < l. \quad (1)$$

[в точках x разрыва функции $f(x)$ сумма ряда равна $\frac{f(x-0) + f(x+0)}{2}$], коэффициенты которого определяются равенствами (4) § 17:

$$a_n = \frac{1}{l} \int_{-l}^l f(t) \cos \frac{n\pi}{l} t dt, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{1}{l} \int_{-l}^l f(t) \sin \frac{n\pi}{l} t dt, \quad n = 1, 2, \dots$$

Примечание. Для того, чтобы отличать x от переменной интегрирования, последнюю обозначили через t .

Выясним, как изменится правая часть равенства (1), если перейти к пределу при $l \rightarrow +\infty$. Для этого сначала преобразуем ряд, записанный в правой части равенства (1). Подставляя в него вместо a_n и b_n их выражения через интегралы, получаем:

$$f(x) = \frac{1}{2l} \int_{-l}^l f(t) dt + \frac{1}{l} \sum_{n=1}^{\infty} \left[\int_{-l}^l f(t) \cos \frac{n\pi}{l} t dt \cdot \cos \frac{n\pi}{l} x + \int_{-l}^l f(t) \sin \frac{n\pi}{l} t dt \cdot \sin \frac{n\pi}{l} x \right]$$

или, что то же,

$$f(x) = \frac{1}{2l} \int_{-l}^l f(t) dt + \frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \frac{n\pi}{l} (t-x) dt. \quad (2)$$

Равенство (2) имеет место для всех x , удовлетворяющих условию $|x| < l$, где l — любое фиксированное число. Пусть теперь x — любая фиксированная точка, тогда, рассматривая в равенстве (2) различные значения l , $l > |x|$, перейдем в нем к пределу при $l \rightarrow +\infty$.

Прежде всего замечаем, что при $l \rightarrow +\infty$ $\lim_{l \rightarrow +\infty} \frac{1}{2l} \int_{-l}^l f(t) dt = 0$.

Действительно, так как интеграл $\int_{-\infty}^{+\infty} |f(t)| dt$ сходится, то его частные интегралы ограничены, в частности, ограничены и интегралы, взя-

¹⁾ См. § 16, достаточные признаки разложимости функции в ряд Фурье.

тые по промежуткам, симметричным относительно точки $x=0$, т. е. существует такое число $M > 0$, что $\int_{-l}^l |f(t)| dt \leq M$ при любом l . Так как

$$\left| \int_{-l}^l f(t) dt \right| \leq \int_{-l}^l |f(t)| dt,$$

то, значит,

$$\left| \int_{-l}^l f(t) dt \right| \leq M$$

и $\lim_{l \rightarrow +\infty} \frac{1}{2l} \left| \int_{-l}^l f(t) dt \right| = 0$ или, что то же, $\lim_{l \rightarrow +\infty} \frac{1}{2l} \int_{-l}^l f(t) dt = 0$.

Таким образом, переходя в равенстве (2) к пределу при $l \rightarrow +\infty$, получаем:

$$f(x) = \lim_{l \rightarrow +\infty} \frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \frac{n\pi}{l} (t-x) dt. \quad (3)$$

Пусть $\frac{n\pi}{l} = u_n$ (последовательность $\{u_n\}$ изображена на рис. 447), тогда $\Delta u_n = u_{n+1} - u_n = \frac{\pi}{l}$ и, следовательно, $\frac{1}{l} = \frac{\Delta u_n}{\pi}$. При $l \rightarrow \infty$, $\Delta u_n \rightarrow 0$ и равенство (3) можно записать в виде:

$$f(x) = \frac{1}{\pi} \lim_{\Delta u_n \rightarrow 0} \sum_{n=1}^{\infty} \Delta u_n \int_{-l}^l f(t) \cos u_n (t-x) dt. \quad (4)$$

Интегралы, входящие под знак суммы, являются значениями функции

$$\Phi(u, x) = \int_{-l}^l f(t) \cos u (t-x) dt$$

в точках последовательности $\{u_n\}$ при фиксированном x . Вследствие этого сумма в правой части равенства (4) внешне, в какой-то степени, напоминает интегральную сумму по переменной u ¹⁾, составленную для промежутка $0 \leq u < +\infty$. Примем без доказательства, что предел суммы, записанной в правой части равенства (4), равен интегралу

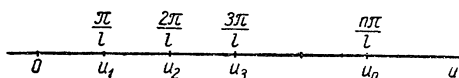


Рис. 447.

$$\int_0^{+\infty} du \int_{-\infty}^{+\infty} f(t) \cos u (t-x) dt.$$

Тогда в любой точке непрерывности функции $f(x)$ будет выполняться равенство:

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} du \int_{-\infty}^{+\infty} f(t) \cos u (t-x) dt \quad (5)$$

(в точке x разрыва функции $f(x)$ вместо $f(x)$ в левой части равенства получится $\frac{f(x-0) + f(x+0)}{2}$).

¹⁾ Напомним, что в интегральной сумме число слагаемых конечно и промежутков, для которого составлена сумма, также конечен.

²⁾ При этом внутренний интеграл $\int_{-\infty}^{+\infty}$ абсолютно сходится, так как модуль под-интегральной функции не больше модуля $f(t)$.

Преобразуем теперь правую часть равенства (5), используя формулу для косинуса разности:

$$f(x) = \int_0^{+\infty} du \left[\frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \cos ut \, dt \cdot \cos ux + \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \sin ut \, dt \cdot \sin ux \right]$$

или, иначе,

$$f(x) = \int_0^{+\infty} [a(u) \cos ux + b(u) \sin ux] \, du, \quad (6)$$

где

$$\left. \begin{aligned} a(u) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \cos ut \, dt, \\ b(u) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \sin ut \, dt. \end{aligned} \right\} \quad (7)$$

Пусть $f(x)$ — любая абсолютно интегрируемая на всей числовой оси функция, непрерывная или имеющая конечное число точек разрыва первого рода на каждом конечном отрезке.

Интегралом Фурье этой функции называется интеграл

$$\int_0^{+\infty} [a(u) \cos ux + b(u) \sin ux] \, du,$$

где $a(u)$ и $b(u)$ определяются равенствами (7), при этом записывают:

$$f(x) \sim \int_0^{+\infty} [a(u) \cos ux + b(u) \sin ux] \, du.$$

Если в каждой точке непрерывности функции $f(x)$ интеграл Фурье этой функции равен значению функции $f(x)$, то знак соответствия $\sim$ заменяется знаком равенства и говорят, что функция $f(x)$ *представима интегралом Фурье*.

Двойным интегралом Фурье абсолютно интегрируемой на промежутке $(-\infty, +\infty)$ функции $f(x)$, непрерывной или имеющей конечное число точек разрыва первого рода на каждом конечном отрезке, называется интеграл

$$\frac{1}{\pi} \int_0^{+\infty} du \int_{-\infty}^{+\infty} f(t) \cos u(t-x) \, dt.$$

Из приведенных рассуждений вытекает следующая теорема: если $f(x)$ — абсолютно интегрируемая на промежутке $(-\infty, +\infty)$ функция, кусочно-гладкая или кусочно-монотонная на каждом отрезке $[-l, l]$, то для этой функции справедливы равенства (6), (5) или, что то же, в этом случае функция $f(x)$ представима соответственно интегралом Фурье и двойным интегралом Фурье.

Интеграл Фурье для четной и нечетной функций. Пусть $f(x)$ — четная функция, удовлетворяющая условиям представимости интегралом Фурье.

Учитывая, что $\int_{-\infty}^{+\infty} = \lim_{l \rightarrow +\infty} \int_{-l}^l$, а также свойство интегралов по симмет-

ричному относительно точки $x=0$ интервалу от четных функций, из равенств (7) получаем:

$$\left. \begin{aligned} a(u) &= \frac{2}{\pi} \int_0^{+\infty} f(t) \cos ut \, dt, \\ b(u) &= 0. \end{aligned} \right\} \quad (8)$$

Таким образом, интеграл Фурье четной функции $f(x)$ запишется так:

$$f(x) = \int_0^{+\infty} a(u) \cos ux \, du, \quad (9)$$

где $a(u)$ определяется равенством (8).

Рассуждая аналогично, получим, что для нечетной функции $f(x)$, удовлетворяющей условиям представимости интегралом Фурье,

$$\left. \begin{aligned} a(u) &= 0, \\ b(u) &= \frac{2}{\pi} \int_0^{+\infty} f(t) \sin ut \, dt, \end{aligned} \right\} \quad (10)$$

и, следовательно, интеграл Фурье нечетной функции имеет вид:

$$f(x) = \int_0^{+\infty} b(u) \sin ux \, du, \quad (11)$$

где $b(u)$ определяется равенством (10).

Заметим, что если функция $f(x)$ задана на промежутке $[0, +\infty)$ и удовлетворяет на этом промежутке условиям представимости интегралом Фурье (т.е. $f(x)$ — абсолютно интегрируемая функция на промежутке $[0, +\infty)$ и кусочно-гладкая или кусочно-монотонная на любом отрезке $[0, l]$), то формула (9) продолжает функцию $f(x)$ четным образом на всю ось OX , а формула (11) дает нечетное продолжение этой функции $f(x)$ на всю числовую ось.

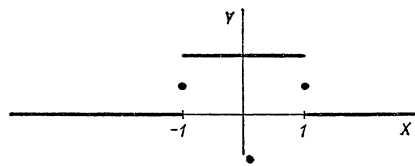


Рис. 448.

Подставляя выражения $a(u)$ и $b(u)$ (формулы (8) и (10)) в равенства (9) и (11), получаем:

$$f(x) = \frac{2}{\pi} \int_0^{+\infty} \cos ux \, du \int_0^{+\infty} f(t) \cos ut \, dt, \quad (12)$$

$$f(x) = \frac{2}{\pi} \int_0^{+\infty} \sin ux \, du \int_0^{+\infty} f(t) \sin ut \, dt. \quad (13)$$

Правые части равенств (12) и (13) для четной и нечетной функций, представимых интегралом Фурье, называются *двойными интегралами Фурье* соответственно *четной и нечетной функций*.

Пример. Представить интегралом Фурье функцию

$$f(x) = \begin{cases} 1, & |x| < 1, \\ \frac{1}{2}, & |x| = 1, \\ 0, & |x| > 1. \end{cases}$$

График этой функции изображен на рис. 448.

Решение. Функция $f(x)$ — нечетная, она удовлетворяет условиям представимости интегралом Фурье. Учитывая это, на основании формулы (11) запишем:

$$f(x) = \int_0^{+\infty} a(u) \cos ux \, du.$$

Применяя формулу (10), определим:

$$\begin{aligned} a(u) &= \frac{2}{\pi} \int_0^{+\infty} f(t) \cos ut \, dt = \frac{2}{\pi} \int_0^1 \cos ut \, dt = \\ &= \frac{2}{\pi} \frac{\cos ut}{u} \Big|_{t=0}^{t=1} = \frac{2}{\pi} \frac{\sin u}{u} \end{aligned}$$

и, следовательно,

$$f(x) = \frac{2}{\pi} \int_0^{+\infty} \frac{\sin u}{u} \cos ux \, du$$

или, что то же,

$$\frac{2}{\pi} \int_0^{+\infty} \frac{\sin u}{u} \cos ux \, du = \begin{cases} 1, & |x| < 1, \\ \frac{1}{2}, & |x| = 1, \\ 0, & |x| > 1. \end{cases}$$

В частности, при $x=0$, получаем:

$$1 = \frac{2}{\pi} \int_0^{+\infty} \frac{\sin u}{u} \, du,$$

т. е.

$$\int_0^{+\infty} \frac{\sin u}{u} \, du = \frac{\pi}{2}.$$

Интеграл Фурье в комплексной форме. Пусть $f(x)$ — функция, представляемая интегралом Фурье, тогда на основании формулы (6)

$$f(x) = \int_0^{+\infty} [a(u) \cos ux + b(u) \sin ux] \, du,$$

при этом по формулам (7) имеем:

$$\begin{aligned} a(u) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \cos ut \, dt, \\ b(u) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \sin ut \, dt. \end{aligned}$$

Преобразуем подинтегральное выражение, используя формулы Эйлера:

$$\begin{aligned} a(u) \cos ux + b(u) \sin ux &= a(u) \frac{e^{iux} + e^{-iux}}{2} + b(u) \frac{e^{iux} - e^{-iux}}{2i} = \\ &= a(u) \frac{e^{iux} + e^{-iux}}{2} - ib(u) \frac{e^{iux} - e^{-iux}}{2} = \\ &= \frac{a(u) - ib(u)}{2} e^{iux} + \frac{a(u) + ib(u)}{2} e^{-iux}. \end{aligned}$$

Если обозначить

$$\left. \begin{aligned} \frac{a(u) - ib(u)}{2} &= c(u), \\ \frac{a(u) + ib(u)}{2} &= c(-u), \end{aligned} \right\} u \geq 0,$$

то подинтегральная функция запишется так:

$$a(u) \cos ux + b(u) \sin ux = c(u) e^{iux} + c(-u) e^{-iux}$$

и, значит,

$$f(x) = \int_0^{+\infty} [a(u) \cos ux + b(u) \sin ux] du = \int_0^{+\infty} [c(u) e^{iux} + c(-u) e^{-iux}] du$$

или, иначе

$$f(x) = \lim_{A \rightarrow +\infty} \int_0^A [c(u) e^{iux} + c(-u) e^{-iux}] du. \quad (14)$$

Преобразуем теперь частные интегралы, входящие в равенство (14):

$$\begin{aligned} \int_0^A [c(u) e^{iux} + c(-u) e^{-iux}] du &= \int_0^A c(u) e^{iux} du + \int_0^A c(-u) e^{-iux} du = \\ &= \int_0^A c(u) e^{iux} du + \int_{-A}^0 c(u) e^{iux} du = \int_{-A}^A c(u) e^{iux} du. \end{aligned}$$

(При этом мы учли, что интеграл $\int_0^A c(-u) e^{-iux} du$ с помощью подстановки

$-u = v$ примет вид $\int_{-A}^0 c(v) e^{ivx} dv$, который, как известно, не зависит от написания переменной интегрирования и поэтому может быть записан в виде $\int_{-A}^0 c(u) e^{iux} du$.)

Учитывая это, равенство (14) запишем следующим образом:

$$f(x) = \lim_{A \rightarrow +\infty} \int_{-A}^A c(u) e^{iux} du$$

или, что то же,

$$f(x) = \int_{-\infty}^{+\infty} c(u) e^{iux} du. \quad (15)$$

Найдем теперь выражение для $c(u)$. Так как

$$\begin{aligned} c(u) &= \frac{a(u) - ib(u)}{2} = \frac{1}{2} \left(\frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \cos ut dt - i \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \sin ut dt \right) = \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) (\cos ut - i \sin ut) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) e^{-iut} dt, \end{aligned}$$

то окончательно

$$c(u) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) e^{-iut} dt. \quad (16)$$

Равенство (16) получено при условии, что $u \geq 0$, нетрудно видеть, что оно сохраняется и при $u < 0$. Действительно,

$$\begin{aligned} c(-u) &= \frac{a(u) + ib(u)}{2} = \frac{1}{2} \left(\frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \cos ut dt + \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \sin ut dt \right) = \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) e^{iut} dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) e^{-i(-u)t} dt. \end{aligned}$$

¹⁾ Здесь несобственный интеграл понимается как $\lim_{A \rightarrow +\infty} \int_{-A}^A$.

Из сравнения выражений для $c(u)$ и $c(-u)$ заключаем, что равенство (16) верно при всех действительных значениях u . Выражение $\int_{-\infty}^{+\infty} c(u) e^{iux} du$ называется *интегралом Фурье в комплексной форме* функции $f(x)$.

Если выражение $c(u)$ подставить в равенство (15), то получим:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{iux} du \int_{-\infty}^{+\infty} f(t) e^{-iut} dt$$

или после внесения e^{iux} под знак внутреннего интеграла:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} f(t) e^{iu(x-t)} dt, \quad (17)$$

где наружный интеграл понимается как $\lim_{A \rightarrow +\infty} \int_{-A}^A$.

Выражение $\frac{1}{2\pi} \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} f(t) e^{iu(x-t)} dt$ называется *двойным интегралом Фурье в комплексной форме* функции $f(x)$.

Преобразование Фурье. Пусть $f(x)$ — любая абсолютно интегрируемая на всей числовой оси функция, непрерывная или имеющая конечное число точек разрыва первого рода на каждом конечном отрезке. *Преобразованием Фурье*¹⁾ такой функции $f(x)$ называется функция

$$F(u) = \int_{-\infty}^{+\infty} f(x) e^{-iux} dx. \quad (18)$$

Если $f(x)$ — функция, представимая интегралом Фурье, то на основании равенства (17):

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(u) e^{iux} du. \quad (19)$$

Равенство (19) называется *формулой обращения преобразования Фурье*, или обратным преобразованием Фурье. Равенством (19) функция $f(x)$ определяется через ее преобразование Фурье.

Таким образом, если рассматривать интегральное уравнение (19), в котором функция $f(x)$ — заданная, а функция $F(u)$ — искомая, то его решение определяется равенством (18).

Аналогично, если имеем интегральное уравнение вида (18), в котором $F(u)$ — заданная функция, а $f(x)$ — искомая, то его решение определяется равенством (19).

¹⁾ Преобразование Фурье функции $f(x)$ называется также *спектральной характеристикой функции $f(x)$* .

*Игнатьева Алла Венедиктовна,
Краснощекова Таисия Ивановна,
Смирнов Виктор Федорович*

КУРС ВЫСШЕЙ МАТЕМАТИКИ

Редактор *Тальский Д. А.*
Технический редактор *Григорчук Л. А.*
Корректор *Шмелева И. С.*

Т-12532. Сдано в набор 13/IV-64 г. Подписано к печати 25/IX-64 г. Формат 70×108¹/₁₆. Объем 42,75 печ. л. 59,85 усл. п. л. 49,37 уч.-изд. л. Изд. № ФМХ/132. Тираж 92000 экз. Зак. 1000. Цена 1 р. 56 к.
Москва, И-51, Неглинная ул., д. 29/14
Издательство «Высшая школа»

Ленинградская типография № 1 «Печатный Двор» имени А. М. Горького
«Главполиграфпрома» Государственного комитета Совета Министров СССР
по печати, Гатчинская, 26.

